5. Do contingent valuation estimates of willingness to pay for non-use environmental goods pass the scope test with adequacy? A review of the evidence from empirical studies in the literature[1]

James Burrows, Rebecca Newman, Jerry Genser, and Jeffrey Plewes[2]

INTRODUCTION

Contingent valuation (CV) is commonly used in environmental economics to estimate non-use values of environmental goods and services.[3] Use values are amenable to direct analysis based on revealed preference data, either from direct approaches such as reviewing evidence of actual

[1] The authors gratefully acknowledge the essential contributions made to this chapter by Ed Leamer, Ken Train, Renée Miller-Mizia, Stamatia Kostakis, Hasat Cakkalkurt, Hiu Man Chan, and Connor Tobin.

[2] Respectively, Vice Chairman, Charles River Associates, Boston; Associate, Charles River Associates, Boston; Associate, Charles River Associates, Boston; Principal, Charles River Associates, Boston.

[3] Contingent valuation (CV) is a survey-based methodology often used to estimate values of non-market resources, such as environmental amenities. The survey may directly ask respondents how much they would be willing to pay (or willing to accept) for an environmental amenity (or to avoid the loss of an environmental amenity). A choice experiment (CE) is an application of the CV method in which respondents are presented with multiple questions with two or more choices in each question and are asked to select the preferred alternative in each choice question. The choice data are then econometrically analyzed to infer WTPs. CV can employ a variety of elicitation techniques, including open ended (in which respondents are asked to specify their WTP without a prompted amount), dichotomous choice (DC), in which respondents are asked whether their willingness to pay is at least an offered cost, double-bounded and multiple-bounded dichotomous choice (in which respondents are asked follow-up valuation questions to narrow the range of their WTPs), payment cards (in which respondents are shown a payment card with suggested values), and bidding games, among others.

purchases, or indirect approaches such as travel cost and hedonic price analysis. CV is likely to be much less accurate for non-use amenities than for use amenities, as respondents asked about non-use amenities have no market experience to guide their thinking, and are unlikely to have ever given thought to assigning monetary values for this type of good.

From its earliest days, the CV method has been scrutinized with respect to whether its results are consistent with the assumptions of rational choice. In this chapter we focus on one of the key tests of rational choice: do estimates of willingness to pay (WTP) derived from CV studies increase as the amount of the good (or the number of goods) increases (i.e., as scope increases), and, if so, are the WTP estimates "adequately" responsive to scope?[4]

A fundamental tenet of consumer utility theory is that utility increases as consumption of most goods increases (i.e., as the scope of consumption increases). For most goods and services, marginal utility generally declines as consumption increases; accordingly, some decline in WTP per unit as scope increases is to be expected for most goods over a reasonable range of costs or prices. A finding in a study that demand for an environmental amenity is not scope sensitive (i.e., the WTP for the amenity does not increase with scope, or increases by an amount that is too small to be credible) can occur for a number of reasons: respondents in fact do not experience increasing utility from increased scope of the amenity,[5] the methodology may be flawed, or respondents receive a "warm glow" from indicating (to themselves and/or the survey administrator) that they are willing to contribute to a worthy cause. If a flawed methodology or warm glow are the cause of scope non-responsiveness, the WTP results from the study are not usable for applications of CV,

[4] We define scope that is in terms of a single argument (e.g., acres of clean beaches) as quantitative scope, and scope that is in terms of multiple arguments (e.g., miles of clean beaches and number of lakes) as categorical scope. In the environmental literature the term "embedding" is often used to refer to instances of low or no scope sensitivity (with perfect embedding indicating that the value of a larger quantity is equal to the value of a smaller quantity).

[5] For example, it is quite possible that there is a positive WTP for wolves up to some minimum population size, but that the respondents on average may have no WTP (or even negative WTP) for additional numbers of wolves above this minimum population size. It is also possible that respondents may experience negative utility from any number of wolves. Boman and Bostedy (1999) and Wilson (2000) find no scope sensitivity for wolves. These wolf results are excluded from our tabulations because it is possible the respondents have no underlying positive marginal utility for wolves. As another example, there may be scope insensitivity for protecting land that gives access to a scenic body of water – WTP for preserving enough land for access may be high, but WTP for additional land above this threshold may be small.

including the allocation of public funds or the determination of damages in a litigation matter.

It would not be surprising to find a few studies in a large sample showing low or no scope effects. However, a finding that a large percentage of CV studies report no or very low scope effects would suggest that most environmental goods are subject to highly diminishing marginal utility, which, prima facie, seems improbable – for example, in the case of *quantitative* scope insensitivity is it really believable that respondents consistently place a high value on preserving one species, one pristine lake, or one pristine forest, and little or no value on additional species, pristine lakes, or pristine forests? In the case of *categorical* scope insensitivity, is it believable that respondents place a high value on preserving one species and little or no value in preserving one species and restoring one lake in addition? While previous reviews of the literature addressed how many studies pass or fail scope tests, none, to our knowledge, also attempted to quantify the adequacy or plausibility of these results over multiple studies.

We seek to review scope tests based on split-sample results ("external" scope tests in which responses of separate, independent groups are used to infer WTP for one scope of the environmental amenity) in CV studies of environmental amenities to determine the proportion that pass or fail. We also quantify the marginal utility implied by scope tests that pass, where applicable. We do not include within-respondent ("internal") scope tests because respondents who are presented with more than one choice task may attempt to appear internally consistent (either to themselves or to the survey administrator) and may also be affected by anchoring effects (to costs or prices that they have already seen in the survey) and/or by other context effects. We do not include scope tests derived from choice experiment (CE) surveys because scope tests using this methodology are, by their nature, internal tests in which the same respondents are making choices involving multiple scopes of environmental amenities. We focus in particular on scope tests involving non-use environmental goods, but we also include scope tests involving use environmental goods. Our review of 111 papers disproves the widely held conclusion that scope tests typically pass, as the majority of tests surveyed here fail. Those that do pass tend to do so with an implausibly low marginal WTP for additional units of the environmental amenity.

Scope insensitivity is often attributed to diminishing marginal utility and satiation. Amiran and Hagen (2010) argue that sharply diminishing marginal utility leading to low scope responsiveness can be explained by "bounded substitution" between environmental goods and market goods, in which survey respondents are willing to make few trades between money and environmental goods. Respondents may be willing to trade money

for the first increment of an environmental amenity, but not additional increments.

Sharply diminishing marginal utility may well be true for particular goods over a particular range of prices. However, the instances in which this can be invoked are likely to be limited. Sharply diminishing marginal utility explains findings of limited scope only if respondents' preferences take a particular form: marginal utility from additional environmental goods must begin to diminish immediately after obtaining the smaller increment of the environmental good. Consider, for example, Araña and León (2008). For respondents with average emotional intensity scales, the authors find a WTP of €15.29 to increase the current length of walking paths from 300 to 330 kilometers (€0.50 for each additional kilometer). However, respondents are willing to pay only slightly more (€16.64) for a larger environmental good that would increase the length of the walking paths from 300 to 400 kilometers (€0.02 per kilometer, for the additional 70 kilometers). These results imply respondents are willing to pay nearly 30 times as much money per kilometer for the first 30 kilometers that are restored than they are for the remaining 70. Diminishing marginal utility rationalizes this finding only if survey respondents become satiated with kilometers soon after restoring the first 30 kilometers. Furthermore, if diminishing marginal utility does explain this result, it must be the case that Araña and León (2008) would have found increased sensitivity to scope if their large environmental good were smaller in size. If, for example, the larger environmental good increased the trail length from 300 to 350 kilometers (instead of 400), the empirical results indicate that the authors would necessarily have found WTPs per kilometer that were more similar across the smaller and larger environmental goods. In other words, findings similar to those of Araña and León (2008) can be rationalized by diminishing marginal utility only if we agree that the survey was accidentally designed not to find evidence of scope. Such accidents should be uncommon. However, as we show later, WTPs that imply sharply diminishing marginal utility are the rule and not the exception for those studies that find some scope responsiveness.

Carson and Mitchell (1989) identified a number of other factors that could also lead to apparent scope insensitivity: (1) part-whole bias, in which respondents confuse the good being offered with a much larger or smaller good;[6] (2) symbolic bias, in which respondents might perceive a

[6] Carson states: "Another problem. . .occurs when the researcher believes one good encompasses another, but respondents find the two goods offered indistinguishable. For example, suppose an ecosystem that provides habitat for five species is at risk. As a scope test, the researcher informs one sample that the habitat will be purchased to protect the five spe-

good as symbolic of a larger good; (3) metric bias, in which a respondent might be defining a good in a different metric than the survey designer; and (4) probability of provision bias, in which the respondent might believe that the larger good has a lower probability of being provided than the smaller good, and therefore bids less for the larger good than he or she otherwise would.

Biases of the types identified by Carson and Mitchell may well exist, and some may be curable by improved survey design. However, these types of biases considerably complicate, if not make impossible, the interpretation of a CV study. If a CV study subject to considerable part-whole bias finds that average WTP of respondents to clean up a mile of soiled beaches is, say, $10, how is the decision maker to know whether respondents are really valuing a much bigger good, such as all the beaches in a broad region? If the answer is that well-designed studies avoid the biases identified by Carson and Mitchell, what objective criteria can the decision maker use to determine whether a particular CV study is sufficiently well designed to be credible?

An alternative explanation of scope insensitivity is "warm glow" (see Andreoni, 1990 and Kahneman and Knetsch, 1992) – that is, respondents purchase "moral satisfaction" when they bid on a good (hypothetically or in actuality). If warm glow is fixed in size (scope insensitive) and large relative to the underlying marginal utility of the good being valued, then the WTP for that good will not increase very much as the quantity of the good increases. If scope insensitivity is caused by warm glow, estimates of WTP from CV surveys may not be informative about underlying WTP, as much of the estimated WTP may be for warm glow rather than for the amenity in question.

The *Exxon Valdez* spill in 1990 spawned a number of articles questioning or defending the CV methodology. In response to the controversy surrounding CV, the National Oceanic and Atmospheric Administration (NOAA) convened a panel of distinguished experts to review whether "the CV technique is capable of providing reliable information about lost existence or other passive-use values."[7]

The Panel examined various criteria of reliability of CVM, including whether CV estimates of WTP exhibit rationality:

cies 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" (Carson, 1997, pp. 128–9).

[7] Arrow et al. (1993), p. 5.

> Rationality in its weakest form requires certain kinds of consistency among choices made by individuals. For instance, if an individual chooses some purchases at a given set of prices and income, then if some prices fall and there are no other changes, the goods that the individual would now buy would make him or her better off. . . Usually, though not always, it is reasonable to suppose that more of something regarded as good is better so long as an individual is not satiated. This is in general translated into a willingness to pay somewhat more for more of a good, as judged by the individual. Also, if marginal or incremental willingness to pay for additional amounts does decline with the amount already available, it is usually not reasonable to assume that it declines very abruptly.[8]

In its review of the literature the Panel identified a number of "maladies" that would render a CV study "unreliable." In particular, citing articles by Kahneman and Knetsch (1992), Desvousges et al. (1992), and Diamond et al. (1993), it observed that evidence supporting embedding had "multiplied" since Kahneman published a well-known Ontario study.[9] The Panel stated that:

> [. . .]average willingness to pay is often substantial for the smallest scenario presented but is then substantially independent of the size of the damage averted. [This is] potentially a very damaging criticism of the method. . . If reported willingness to pay accurately reflected actual willingness to pay, then, under the "warm glow" interpretation, willingness to pay might well exceed compensation required because the former contains an element of self-approbation. It might be real but not properly compensable.[10]

[8] Ibid., pp. 10–11.

[9] Kahneman (1986) presents a chart from a split-sample telephone survey showing expressed WTP for three different scopes of lake clean-up (Muskoka only, Haliburton only, and all of Ontario). Kahneman states that "[t]he demand functions for the three cleanup operations are strikingly similar" and that "[t]he results indicate that people seem to be willing to pay almost as much to clean up one region or any other, and almost as much for any one region as for all Ontario together. We know from other surveys that these responses do not reflect expectations of personal enjoyment from the cleanup, since Toronto residents are willing to pay substantial amounts to clean up the lakes of British Columbia!. . . Because the questions all elicit symbolic expressions of the same attitude, there is not much difference between the numbers that are attached to a single region and to all of Ontario" (pp. 191–2). Carson (1997) observes that the Kahneman results demonstrate some limited scope sensitivity, as the WTP for clean-up of all Ontario lakes appears from the graph to exceed the WTPs for the smaller goods by about 50%. Desvousges et al. (1992) conduct a survey of WTP to prevent 2,000, 20,000, or 200,000 migratory waterfowl from dying in waste-oil holding ponds and report that they do not find scope sensitivity. Diamond et al. (1993) report on a survey of WTP to prevent various wilderness areas from being developed in which they generally do not find scope sensitivity.

[10] Arrow et al. (1993), pp. 26–7.

The Panel concluded that the findings of a survey would be "unreliable" if the survey exhibited "Inadequate responsiveness to the scope of the environmental insult."[11,12]

In a later document, four of the authors of the NOAA report clarified what they meant by the use of the word "inadequate":

> Had the panel thought that something as straightforward as statistical measurability were the proper way to define sensitivity, then we would (or should) have opted for language to that effect. A better word than "adequate" would have been "plausible": A survey instrument is judged unreliable if it yields estimates which are implausibly unresponsive to the scope of the insult. This, of course, is a judgment call, and cannot be tested in a context-free manner.[13]

We review the evidence on both the extent to which CV studies demonstrate significant sensitivity to scope and the plausibility of scope results that are measured. In accordance with the NOAA Panel's recommendation, any judgment about whether CV inferences about WTP are reliable needs to consider both statistical significance and plausibility of reported scope effects.

Diamond et al. (1993) proposed an "adding-up" test for the reliability of CV that does not rely on untestable assumptions about the structure of respondents' preferences with respect to diminishing marginal utility and bounded substitution.[14] Underlying the adding-up test is the following idea: if stated preference survey responses reflect well-formed stable preferences, then alternative measures of these preferences should yield similar effects. If alternative measures of the same preferences do not yield similar results, then the measures must be measuring something other than preferences.

Specifically, suppose we are interested in measuring willingness to pay for two environmental goods, A and B, and for the combination of the two goods, A + B = C. While a conventional scope analysis tests whether WTP (C) is ≥ WTP (A) or WTP (C) ≥ WTP (B), in an adding-up test the analyst

[11] Ibid., p. 37.

[12] Carson (1994) states: "As used by the Ohio court and in the NOAA Panel report, the reliability of a measure is the degree to which it measures the theoretical construct under investigation. However, in the empirical social sciences, this preceding definition pertains to *validity*, whereas reliability is defined as the extent to which the variance of the measure is not due to random sources and systematic sources of error. . .we. . .use the term *reliability* to refer to the degree to which CV surveys measure the theoretical construct under investigation" (footnote 11, p. 8; emphasis in original). Note that Carson, as is typical of the CV literature, is in effect focusing exclusively on statistical significance without regard to adequacy of scope.

[13] Arrow et al. (1994).

[14] See also Diamond (1996).

tests whether WTP (A) + WTP (B|A) = WTP (C). If respondents' answers reflect well-formed stable preferences over the costs and benefits associated with A and B, then it must be true that the sum of WTP (A) and WTP (B|A) approximately equals WTP (A + B). This equality must hold even if diminishing marginal utility or bounded substitution cause WTP (A + B) to be only slightly larger than WTP (A). However, if measures of WTP are contaminated by phenomena such as "warm glow," this equality will not hold. Upward bias in WTP caused by "warm glow" will cause the sum of WTP (A) and WTP (B|A) to exceed WTP (A + B). Thus, if WTP (A + B) is statistically different from the sum of WTP (A) and WTP (B|A), we can reject the hypothesis that survey responses are a measure of well-formed stable preferences.

The adding-up test and the scope test are related but are not the same. A survey can have zero or implausibly low scope and still pass an adding-up test. On the other hand, if warm glow is an important element of measured WTP (and if warm glow is substantially exhausted after the first "purchase"), a survey will fail both an adding-up test and a scope test.

Critics of the adding-up test object that the scenario presented to the respondents is implausible and difficult to describe: "[I]t is an ex-post counterfactual scenario. Respondents must be convinced that a currently nonexistent government program has been funded and implemented and that their budget has been reduced by the cost."[15] Although this is true, some published CV studies posit hypothetical scenarios in which the government has already completed some environmental investments and the respondent is now being asked to state WTP for more improvements (or ask the respondent to assume some base scenario of environmental characteristics about which the respondent has no prior knowledge).[16] The only manner in which the scenario in an adding-up test is different from CV studies like those of Rollins and Lyke (1998) and Whitehead et al. (2009) is asking the respondent to assume that he or she has already made a payment for A and that A has already been provided when he or she is asked to value B. However, even if the respondent does not take the cost of having paid for A into account in valuing B, the effect on estimated WTP for B is likely to be trivial – in most cases, the cost of B will be a tiny

[15] Whitehead (2016), p. 19.

[16] For example, Rollins and Lyke (1998) inform respondents that the government has set aside 29 of 39 natural regions as national parks and is now considering creating additional national parks; Whitehead et al. (2009) inform respondents that the government has set aside 9,000 acres of Saginaw Bay to be protected and is considering purchasing and protecting additional acreages.

percentage of wealth or income, so the effect of the cost of A on WTP for B should be similarly insignificant.[17]

The adding-up test has also been criticized because it increases survey costs.[18] An adding-up test requires at least three split samples, while a traditional scope test can be performed with as few as two split samples. However, many studies with scope tests contain more than two split samples, so cost can scarcely be the reason for avoiding an adding-up test. For example, a number of studies testing part-whole biases use multiple split samples that survey WTP for various goods separately and combined.[19]

In spite of the rigor of the Diamond et al. (1993) adding-up test, it has only been implemented in a handful of studies: Samples and Hollyer (1990) report adding-up tests that fail; Diamond et al. (1993) report adding-up tests that fail; Chapman et al. (2009) conduct a survey that would permit an adding-up test, but they do not perform an adding-up test (although they report a traditional scope test that exhibits scope sensitivity);[20] Desvousges and Matthews (2012) use the Chapman data to perform an adding-up test that fails; and Desvousges et al. (2016) conduct a new survey similar to the Chapman survey with additional scope variations and report adding-up tests that fail. Other studies that have examined part-whole and sequencing biases have used approaches that have most but not all of the elements of an adding-up test.[21]

[17] Hausman and Newey (2016) develop bounds that take account of the share of income spent on a good and its income derivative. Their results demonstrate that for the typical cost of contingent valuation goods, which are typically in the range of $10 to $200, the bounds are almost identical for consumer surplus.

[18] Whitehead (2016), p. 9.

[19] Furthermore, there is hardly any feature of a properly conducted CV study that is not subject to a cost–quality trade-off.

[20] Because these data fail the adding-up test, we report this paper as a fail in our tabulations.

[21] For example, Stevens et al. (1995) perform split-sample surveys that are close to a full adding-up test: they ask respondents in three split samples to value two rare plants, flood protection, and two rare plants plus flood protection. Respondents for the rare plants alone and flood protection alone scenarios are not told to assume that they have already purchased the other amenity, so in this respect the Stevens et al. study does not satisfy all the conditions of the adding-up test (but this should not have a substantial effect on the results). The WTPs for the two rare plants alone plus flood protection alone are more than two rare plants and flood protection combined, indicating that this modified adding-up test fails. Warm glow is a possible explanation for this result. Substitution between rare plants and food protection could also explain the result, but it seems unlikely that rare plants and food protection are partial substitutes in any other sense than that contributing to them provides warm glow. Samples and Hollyer (1990) derive WTPs from three split samples for humpback whales alone, monk seals alone, and whales and seals directly. They report that the WTPs of whales alone + seals alone is greater than whales and seals combined. In this example, it is possible that whales and seals are partial substitutes for each other, although it still seems more likely is that the substitution effect is derived from warm glow even here.

Carson (1997) presents the first systematic review of scope studies.[22,23] This review focuses on the extent to which CV studies find statistically significant scope effects, but does not address the issue of whether reported scope results are plausible or adequate – the criterion stressed by the NOAA Panel. Carson presents a table enumerating 31 studies appearing between 1986 and 1997 that he states: "contain a rejection of the scope insensitivity hypothesis at $p < 0.10$. Most of the studies contain a rejection at $p < 0.05$ and many contain rejections at less than $p < 0.001$."[24] A casual reading of this statement would suggest that the 31 studies in the list all pass scope, but the precise reading is that each paper in the list includes at least one test passing a scope test even if other tests reported in the paper fail. For example, the Carson list includes the Diamond et al. (1993) paper, which reports two scope tests that pass and 43 scope tests that fail.

We differ from Carson with respect to nine of the studies he includes in his table: four studies do not include an external scope test,[25] three studies report mixed results,[26] and in two studies the scope tests fail.[27] In addition, we identify 34 additional scope studies published prior to 1997 involving environmental amenities studies that are not included in the Carson study; as shown in Table 1, seven of these additional studies pass (P) scope, 21 fail (F) scope, and six have mixed (M) results. With fractional allocations for studies reporting multiple scope tests, 9.7 pass scope and 24.3 fail scope. Carson also ignores the key issue of reliability of CV highlighted by the NOAA Panel – namely, whether CV studies have scope results that are adequate. This omission is not unique to Carson; CV researchers often ignore this issue entirely.

The only other systematic review of scope tests in CV studies besides Carson (1997) is Desvousges et al. (2012) (DMT). DMT include a table identifying 109 CV studies in which scope tests are reported (or contain information that permits a scope test even if not reported in the study). They report that more of these 109 studies pass scope than fail: 40 pass (P)

[22] Brown and Duffield (1995) present a table summarizing 14 CV studies with scope results, but this table includes both internal and external scope tests and omitted numerous other studies.

[23] Carson (1997) also includes a version of the Schkade and Payne (1993) bird study data without outliers. Because the standard errors overlap for the outlier removal across each size threshold and model, we tabulate this paper as a fail in our database. This is weighted as 0.5 with the original Schkade and Payne (1993) paper, which we list as 0.5 pass, 0.5 fail.

[24] Carson (1997), Table VI, pp. 143–6.

[25] Duffield and Neher (1991), Whitehead (1992), Wu (1993), and Boyle et al. (1993).

[26] Magnussen (1992), Loomis et al. (1993), Schkade and Payne (1994). These articles are not technically misrepresented in Table VI, as each contains at least one test that passes scope in addition to scope tests that fail.

[27] Navrud (1989); Diamond et al. (1993).

Table 1 Carson (1997) comparison

	Authors	Pub Year	Commodity	External Test Pass/Fail/Mixed	Included in Carson (1997)?	Match Carson's Assessment?
[1]	Brookshire et al.	1983	Benefit of hunting big horn sheep or grizzly bears in five or 15 years	F	No	
[2]	Rahmatian	1985	Visibility in Grand Canyon National Park	F	No	
[3]	Kahneman	1986	Clean-up to preserve fishing in Muskoka, Haliburton, and all Ontario	M	No	
[4]	Bergstrom and Stoll	1987	Farmland protection in Greenville County, South Carolina	F	No	
[5]	Navrud	1989	Reduced sulfur depositions to protect freshwater fish populations	F	Yes	No
[6]	Carson et al.	1990	Air quality in Cincinnati	P	No	
[7]	Samples and Hollyer	1990	Preserving Humpback whales and Hawaiian monk seals in Hawaii	F	No	
[8]	Duffield and Neher	1991		N/A	Yes	No
[9]	Gilbert et al.	1991	Lye Brook Wilderness Area in southwestern Vermont	F	No	
[10]	Desvousges et al.	1992	Preventing waterfowl deaths in the Central Flyway	F	No	
[11]	Desvousges et al.	1992	Preventing environmental damage from oil spills	F	No	
[12]	Jakus	1992	Gypsy moth protection	P	Yes	Yes

[13]	Kahneman and Knetsch	1992	Preserving wilderness, protecting wildlife, providing parks, preparing for disasters, controlling air pollution, insuring water quality, routine treatment of industrial wastes; 12 embedding pairs	M	No	
[14]	Magnussen	1992	Reduce Norwegian nutrient leaching to the North Sea	M	Yes	No
[15]	Rowe et al.	1992	Oil spill clean-up and prevention programs	P	Yes	Yes
[16]	Whitehead	1992		N/A	Yes	No
[17]	Boyle et al.	1993		N/A	Yes	No
[18]	Carson	1993	National versus regional water quality	P	Yes	Yes
[19]	Diamond et al.	1993	Marshall wilderness areas	M	Yes	No
[20]	Fischhoff et al.	1993	River clean-up in Pittsburgh area	F	No	
[21]	Loomis et al.	1993	Protections of forests in Southeast Australia	M	Yes	No
[22]	Greg	1993	Preservation of the Selway-Bitterroot Wilderness	F	No	
[23]	Tanguay et al.	1993	Maintaining woodland caribou population in northwestern Saskatchewan	F	No	
[24]	Wu	1993		N/A	Yes	No
[25]	Bowker and Didychuk	1994	Preserve units of Moncton area farmland	P	Yes	Yes
[26]	Boyle et al.	1994	Preventing waterfowl deaths in the Central Flyway	F	No	
[27]	Carson et al.	1994a	Preservation of Australia's Kakadu conservation zone from mining activity	P	Yes	Yes
[28]	Carson et al.	1994b	Program to accelerate natural restoration of injured resources due to PCB and DDT contamination in Southern California Bight	P	Yes	Yes
[29]	Gerrans	1994	Preservation of the Jandakot wetlands	F	No	

Table 1 (continued)

	Authors	Pub Year	Commodity	External Test Pass/ Fail/Mixed	Included in Carson (1997)?	Match Carson's Assessment?
[30]	Welsh et al.	1994	Glen Canyon Dam downstream recreation, hydropower, and passive-use values	P	Yes	Yes
[31]	Kahneman and Ritov	1994	Headline method – multiple	F	No	
[32]	Krieger	1994	Anglers' WTP for changes in Michigan's public health advisory	P	Yes	Yes
[33]	McFadden	1994	Preservation of the Selway-Bitterroot Wilderness in northern Idaho	F	No	
[34]	Schkade and Payne	1994	Preserving migratory waterfowl in the Central Flyway of the USA (the bird study)	M	Yes	No
[35]	Binger et al.	1995	Preservation of the Selway-Bitterroot Wilderness and 57 federal wilderness areas	P	No	
[36]	Brown and Duffield	1995	Protect instream flow in either one or five Montana rivers	M	No	
[37]	Brown et al.	1995	Preservation of natural areas in Fort Collins	F	No	
[38]	Carson and Mitchell	1995	Preventing injuries from large open pit mine in Kakadu Conservation Zone	P	No	
[39]	Fredman	1995	Protecting the white-backed woodpecker in Sweden	F	No	
[40]	Schultze et al.	1995	Upper Clark Fork River, Montana, River Basin Restoration	F	No	
[41]	Stevens et al.	1995	Preserving wetlands in New England	P	No	
[42]	Welsh et al.	1995	Preserving resources of Colorado River downstream of Glen Canyon Dam	M	No	

[43]	Berrens et al.	1996	Protecting minimum instream flows in New Mexico: silvery minnow versus 11 threatened species	P	No	
[44]	Choe et al.	1996	Improving water quality of rivers and sea in Davao, Philippines	M	No	
[45]	Hoevenagel	1996	Greenhouse effect, depletion of ozone layer, deforestation, acid rain, surface water pollution, animal manure problem	P	Yes	Yes
[46]	Macmillan et al.	1996	Recovery/damage scenarios from reduced acid rain deposition	F	No	
[47]	Carson	1997	Preventing birds from being killed	F	No	
[48]	Huang et al.	1997	Quality improvement of Pamlico and Albemarle sounds (recreation areas) in North Carolina	M	No	
[49]	Loomis and Ekstrand	1997	Preserving MSO or 62 endangered species	F	No	
[50]	Mullarkey	1997	Wisconsin wetlands	P	No	
[51]	Ready et al.	1997	Preventing a decrease in the number of horse farms in Kentucky	P	Yes	Yes
[52]	Smith et al.	1997	Controlling marine debris on beaches and coastal areas in New Jersey and North Carolina	P	Yes	Yes
[53]	Stevens et al.	1997	Movie passes and restoration of Atlantic salmon	P	No	
[54]	White et al.	1997	Conservation of two otter species	F	No	

Table 1 (continued)

Notes:

[5] Navrud (1989) reports the WTP results for two split samples in which respondents are asked WTP for 30, 50, and 70% reductions in sulfur emissions related to corresponding increments in fish populations, with subsamples for payment cards vs bidding game elicitation techniques. Navrud states "When looking at the mean of subsamples using bidding games the WTP appears to increase (but at a decreasing rate) with increasing improvements in the fish populations and for all affected goods. . . However, the observed tendency is not statistically significant, and the same tendency cannot be observed for the subsamples using payment cards. This indicates that the respondents have difficulties in perceiving the differences between the environmental improvements" (p. 81). We report this study as failing scope – with bidding games the results are not significant and for payment cards no scope sensitivity is found.

[8] Duffield and Neher's (1991) study of economic value of waterfowl hunting reports split samples in which respondents are asked for WTP (in trip costs) for waterfowl hunting assuming they saw twice as many (or half as many) waterfowl in their prior trip, but this question occurs after they are asked for their WTP for their last trip – so the results reported are an internal scope test. We record this study as not providing an external scope test.

[14] Magnussen (1992) reports WTP results for four split samples. The first questions are different for each sample (WTPs in parentheses); from largest good to smallest they are: (1) a package of four broad environmental goods (3,054 kroner), (2) a package of four pollution abatement methods (3,366 kroner), (3) a North Sea program (1943 kroner), and (4) improvement in local water quality from class 3 to class 2 (1,125 kroner). WTP for good 2 is higher than for good 1, but this difference is not significant. WTP for good 2 is significantly higher than for good 3. WTP for good 3 is higher than for good 4, but it is only significant at the 10% level. Respondents are also asked subsequent questions (for goods 2 and 3 for the respondents to good 1, for good 3 for respondents to good 2, and for good 5 [improvement from level 2] for respondents to good 2). These are all internal tests and are excluded from our survey (three of these internal tests pass and one fails). The author concludes that "the study did not support that people state the same value in CV surveys whatever amenity is valued. But for very inclusive packages of environmental improvements and 'loose' descriptions of the payment vehicle, this seemed to be the case. . . The small differences in WTP for different marginal water quality improvements question whether people are able to value different marginal changes" (p. 218). We classify this study as providing mixed results with respect to scope sensitivity.

[16] Whitehead's (1992) study of WTP to preserve loggerhead sea turtle species. The respondents are asked to provide their assessments of the probability that the loggerhead sea turtle will become extinct in the next 25 years and then are asked for their WTP for a policy under which the loggerhead sea turtle will definitely not become extinct within the next 25 years. There is no split sample, so there is no external scope test in this study. Nor is there an internal scope test. However, the variable "extinction risk probability change" is significant in both a linear and a non-linear regression of WTP on a number of explanatory variables. This may appear to be a scope test, but the "extinction risk probability" is supplied by the respondents, not the survey administrator. We classify this study as not containing either an internal or external scope test.

[19] Diamond et al. (1993). Carson is correct that this study includes "a rejection of the scope insensitivity hypothesis"; the study reports two scope tests that pass and 43 that fail. Carson does additional analysis on Hypothesis 1 of the Diamond article and reports that the acreage variable is significant in a regression of WTP on acres. We therefore report this is as a pass with a 50% weight, offsetting the Diamond et al. (1993) test of fail with a 50% weight for this particular hypothesis. This changes the pass/fail count to 2.5/42.5. Diamond et al. report a scope test failure derived from Binger et al. (1995) using a similar survey comparing WTP for one part to WTP for 57 parks. We do not include this scope test failure in the Diamond et al. study because it is reported separately in our tabulations as a Binger et al. study. We report the Diamond et al. study as reporting mixed results with respect to scope sensitivity.

[21] Loomis et al.'s (1993) study of WTP for protection of forests in all of southeastern Australia and two smaller portions of that area. The answers to the first questions of each sample allow two scope tests (for East Gippsland vs all Southeast Australia and for the Errinundra Plateau of East Gippsland vs East Gippsland). The first scope test fails and the second passes. The second and third questions for the first sample and the second question for the second sample permit internal scope tests (which pass), but these are not included in our review of external scope tests. We record this study as reporting mixed effects with respect to scope sensitivity.

[24] Wu's (1993) study of value of three different environmental amenities relating to Ohio's Big Darby Creek. Three split samples were asked to value one policy first and then combinations of two and three policies. Only the first questions would permit an external scope test. The WTPs found are $42.80 for enhanced biological diversity, $35.18 for improved streambed visibility, and $19.18 for increasing hiking trails mileage. As these amenities are not nested, this study does not report an external scope test.

[26] Boyle et al.'s (1993) study of the economic value of Grand Canyon white-water boating. Two split samples were asked to value seven water flow scenarios; one sample was asked values for flows from low to high and the other was asked values for flows from high to low (i.e., each group was asked the same questions but in reverse sequence from the other). The authors found a question-ordering for commercial passengers (at the 10% level) but not for private passengers. This study collected data that could have been used for an external scope test as defined here, but no results of any such test are reported in the article. We categorize this study as N/A.

[34] Schkade and Payne's (1994) study of economic value of migratory waterfowl, reports results for three split samples for 2,000, 20,000, and 200,000 bird deaths. The raw WTPs appear mixed (mean WTP for 200,000 birds is greater than for 20,000, but mean WTP for 20,000 is less than mean WTP for 2,000 birds). Carson states that as a result of Schkade and Payne's "small sample sizes, they lack almost any statistical power because the models estimated do not incorporate information that the treatments. . .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. The other problem with the data is that much of the value is driven by a small number of very large outliers. . .dropping out just the two highest observations eliminates the apparent violation of the economic restriction that WTPs increase monotonically with increases in the level of the good. . .dropping the next four largest observations does not change the relative rankings of WTP for the three treatments" (p. 136). We record Schkade and Payne as mixed but also report Carson's re-analysis of the data as a separate study showing a fail. Each of these studies is assigned a 50% weight in our tabulations, as they use the same underlying data.

a scope test, 17 fail (F) a scope test, 47 have mixed (M) results (i.e., report multiple tests, some of which are passes and some of which are fails), and five are listed as "NR" or not reported. However, the DMT tabulation includes both internal and external scope tests. Although not reported in the published article, these results are largely consistent after exclusion of internal tests in the DMT tabulation. Of 71 studies containing external, non-use or mixed non-use and use environmental scope tests, 25 pass, 13 fail, and 33 report mixed results. As in the case in the Carson paper, the DMT paper focuses only on whether CV studies report statistically significant scope effects, not on whether any scope effects reported found are adequate.

Our results differ from the 2012 DMT results, although not by the same extent as our differences with respect to the 1997 Carson results. For the same studies that are included in the DMT survey, as Table 2 shows we report 29 passes (39%), 18 mixed results (24%), and 27 fails (36%) out of a total count of 74 (including results for three studies that DMT include in their table but for which they do not report scope tests). After assigning fractional passes and fails for the mixed scope tests, our count is 37.25 pass; 36.75 fail. The principal difference between the DMT results and the results we present later in this chapter relates to the classification of scope tests as mixed – we report many of these tests as fails or fractionally allocate them.

In addition to the Carson and DMT surveys of scope tests in individual studies, a number of meta-studies have been published that draw inferences about scope effects from WTPs reported across different studies in particular applied areas. Ojea and Loureiro (2011) present a table summarizing the results of 14 meta-studies with respect to scope. They conclude that eight of the studies find positive sensitivity to scope and that six find no or negative scope sensitivity (three studies find insensitivity to scope, and three studies find negative sensitivity to scope). None of the scope meta-studies deals with the issue of adequacy of scope, although a few report data on the quantitative extent of scope.

Table 3 presents an updated version of the Ojea and Loureiro survey. We have included several additional studies that focus on environmental non-use or mixed non-use and use amenities.[28] Similarly to Ojea and Loureiro we excluded meta-studies that focus exclusively on recreational

[28] We have not included Loomis and White (1996) in Table 3, as Richardson and Loomis (2009) present an updated version of the earlier paper, using a very similar methodology and many of the same studies as in the earlier study. We have also replaced Boyle et al. (1994) with Poe et al. (2001), which is an update of Boyle et al. (1994).

Table 2 DMT (2012) comparison (external tests only)

DMT #	Author	Pub Year	Commodity	Results (Pass/Fail/Mixed)	DMT Results (Pass/Fail/Mixed)
DMT_1	Ahearn et al.	2006	Protecting grassland birds in the Central Plains region	M	M
DMT_2	Alvarez-Farzio et al.	1999	Preserving two environmentally sensitive areas in Scotland	P	P
DMT_3	Araña and León	2008	Rehabilitating walking paths in Gran Canaria, Spain	M	M
DMT_5	Banzhaf et al.	2006	Acid rain: quality of water and fish populations, bird species, and tree species (600 lakes versus 900 lakes)	F	P
DMT_8	Bateman et al.	2004	Open access lake improvements located in grounds of University of East Anglia	M	M
DMT_10	Bennett et al.	1998	Dryland salinity in Upper South East region of South Australia	P	M
DMT_12	Berrens et al.	2000	Protecting minimum instream flows in New Mexico: silvery minnow versus 11 threatened species	P	P
DMT_13	Berrens et al.	1996	Protecting minimum instream flows in New Mexico: silvery minnow versus 11 threatened species	P	P
DMT_14	Binger et al.	1995	Preservation of the Selway-Bitterroot Wilderness and 57 federal wilderness areas	P	P
DMT_15	Bliem and Getzner	2008	Ecological restoration in Danube river basin	F	F
DMT_16	Blomquist and Whitehead	1998	Preserving wetland areas in Kentucky	P	M

Table 2 (continued)

DMT #	Author	Pub Year	Commodity	Results (Pass/Fail/Mixed)	DMT Results (Pass/Fail/Mixed)
DMT_17	Bowker and Didychuk	1994	Preserve units of Moncton area farmland	P	P
DMT_18	Boyle et al.	1994	Preventing waterfowl deaths in the Central Flyway	F	F
DMT_20	Brookshire et al.	1983	Benefit of hunting big horn sheep or grizzly bears in five or 15 years	F	M
DMT_21	Brown and Duffield	1995	Protect instream flow in either one or five Montana rivers	M	M
DMT_22	Brown et al.	1995	Preservation of natural areas in Fort Collins	F	M
DMT_23	Carson	1997	Preventing birds from being killed	F	P
DMT_25	Carson et al.	1994b	Program to accelerate natural restoration of injured resources due to PCB and DDT contamination in Southern California Bight	P	P
DMT_26	Carson et al.	1994a	Preservation of Australia's Kakadu conservation zone from mining activity	P	M
DMT_27	Carson et al.	1990	Air quality in Cincinnati	P	NR
DMT_28	Chapman et al.	2009	Alum treatments to prevent algae growth in the Illinois River system and Tenkiller Lake	F	P
DMT_31	Choe et al.	1996	Improving water quality of rivers and sea in Davao, Philippines	M	M
DMT_32	Christie	2001	Recreation opportunities, Grampian region, Scotland	M	M
DMT_35	Day and Mourato	1998	Maintaining river water quality in Beijing	P	P
DMT_36	Desvousges et al.	1992	Preventing waterfowl deaths in the Central Flyway	F	F

	DMT_37	Diamond et al.	1993	Preservation of Selway, Washakie, and Bob Marshall wilderness areas	F	M
	DMT_40	Dupont	2003	Improvements to three recreational activities (swimming, fishing, and boating in Hamilton Harbor, Ontario, Canada)	M	M
	DMT_41	Eom and Larson	2006	Water quality improvement in the Man Kyoung River in South Korea	P	P
	DMT_42	Fischhoff et al.	1993	River clean-up in Pittsburgh area	F	F
	DMT_43	Gerrans	1994	Preservation of the Jandakot wetlands	F	F
	DMT_44	Giraud et al.	1999	Mexican spotted owl and 62 regional threatened & endangered species	P	M
	DMT_45	Goodman et al.	1998	Non-use value of natural coastal environment in England	F	M
	DMT_47	Hanemann	2005	Saving at-risk birds (the bird study revisited)	P	P
	DMT_51	Heberlein et al.	2005	(1) Water quality, (2) Wisconsin's wild wolf population, (3) Indian spearfishing, (4) biodiversity	P	M
	DMT_52	Hite et al.	2002	Subsidizing variable-rate technology to reduce polluted river runoff for the Mississippi River basin (in Mississippi)	P	P
	DMT_53	Hoevenagel	1996	Greenhouse effect, depletion of ozone layer, deforestation, acid rain, surface water pollution, animal manure problem	P	P
	DMT_55	Huang et al.	1997	Quality improvement of Pamlico and Albemarle sounds (recreation areas) in North Carolina	M	P
	DMT_56	Kahneman	1986	Clean-up to preserve fishing in Muskoka, Haliburton, and all Ontario	M	NR
	DMT_57	Kahneman and Ritov	1994	Headline method – multiple	F	M

Table 2 (continued)

DMT #	Author	Pub Year	Commodity	Results (Pass/Fail/Mixed)	DMT Results (Pass/Fail/Mixed)
DMT_58	Kahneman and Knetsch	1992	Preserving wilderness, protecting wildlife, providing parks, preparing for disasters, controlling air pollution, insuring water quality, routine treatment of industrial wastes; 12 embedding pairs	M	M
DMT_60	Krieger	1994	Angler's WTP for changes in Michigan's public health advisory	P	P
DMT_62	Loomis and Ekstrand	1997	Preserving MSO or 62 endangered species	F	P
DMT_66	Loomis et al.	1993	Protections of forests in Southeast Australia	M	M
DMT_67	Macmillan and Duff	1998	Native woodland restoration in the UK	P	P
DMT_68	Macmillan et al.	1996	Recovery/damage scenarios from reduced acid rain deposition	F	M
DMT_69	Magnussen	1992	Reduce Norwegian nutrient leaching to the North Sea	M	M
DMT_71	McDaniels et al.	2003	Benefits of fisheries enhancement on rivers in British Columbia	P	P
DMT_72	McFadden	1994	Preservation of the Selway-Bitterroot Wilderness in Northern Idaho	F	F
DMT_73	Navrud	1989	Reduced sulfur depositions to protect freshwater fish populations	F	M
DMT_75	Nunes and Schokkaert	2003	Protection of wilderness and recreation areas in Alentejo Natural Park, Portugal	M	M
DMT_77	Poe et al.	2005	Mexican spotted owl and 62 regional threatened and endangered species	M	P

DMT_79	Pouta	2005	Forest regeneration cutting policy in Finland	F	M
DMT_80	Powe and Bateman	2004	Protect wetlands (nested area versus total) in Broadland, Eastern England	P	M
DMT_81	Ready et al.	1997	Preventing a decrease in the number of horse farms in Kentucky	P	P
DMT_84	Rollins and Lyke	1998	Creating parks in Canada's Northwest Territories	F	M
DMT_86	Rowe et al.	1992	Oil spill clean-up and prevention programs	P	M
DMT_87	Samples and Hollyer	1990	Preserving humpback whales and Hawaiian monk seals in Hawaii	F	F
DMT_88	Schkade and Payne	1994	Preserving migratory waterfowl in the Central Flyway of the USA (the bird study)	M	F
DMT_90	Smith et al.	2005	Improvements of eastern regional haze	F	M
DMT_91	Smith et al.	1997	Controlling marine debris on beaches and coastal areas in New Jersey and North Carolina	P	P
DMT_92	Stanley	2005	Riverside fairy shrimp versus all local endangered species	P	P
DMT_93	Stevens et al.	1997	Movie passes and restoration of Atlantic salmon	P	F
DMT_94	Stevens et al.	1995	Preserving wetlands in New England	P	P
DMT_95	Streever et al.	1998	Wetland conservation in New South Wales, Australia	F	F
DMT_96	Svedsäter	2000	Rain forests in South America, endangered wild animals, air pollution in central London, global warming	F	F
DMT_97	Tanguay et al.	1993	Maintaining woodland caribou population in Northwestern Saskatchewan	F	NR

Table 2 (continued)

DMT #	Author	Pub Year	Commodity	Results (Pass/Fail/Mixed)	DMT Results (Pass/Fail/Mixed)
DMT_98	Veisten et al.	2004a	Endangered species preservation in Norwegian forests	M	M
DMT_100	Welsh et al.	1995	Preserving resources of Colorado River downstream of Glen Canyon Dam	M	M
DMT_101	White et al.	1997	Conservation of two otter species	F	M
DMT_103	Whitehead and Cherry	2007	Green energy program in North Carolina (yielding improved air quality in western North Carolina mountains)	M	M
DMT_104	Whitehead and Finney	2003	Preserving submerged marine cultural resources (historic shipwrecks)	F	F
DMT_105	Whitehead et al.	2009	Purchase and managing additional acres of coastal marshes in Saginaw Bay	F	F
DMT_106	Whitehead et al.	1998	Quality improvement of Pamlico and Albemarle sounds (recreation areas) in North Carolina	P	P
DMT_108	Wilson	2000	Wolves, Chippewa Indian spearfishing, biological diversity, and water quality in Wisconsin	P	M

Table 3 Ojea and Loureiro (2011) comparison

Study	Year	Commodity	Region	Method	Studies	Obs.	Size Measure	Scope
Non-use								
Barrio and Loureiro	2010	Forests	Worldwide	CV	35	101	Hectares	Insensitive
Brouwer et al.	1999	Wetlands	North America, Europe	CV	30	92	% size of wetland with respect to all country wetlands	Insensitive
Ghermandi et al.	2008	Wetlands	Worldwide	CV, HP, TC, RC, PF, MP, CE	167	385	Hectares	Positive
Hjerpi et al.	2015	Forests and freshwater	Europe, USA, Canada	CV, CR, CE	22	127	Low, High	Positive
Johnston and Duke	2009	Farmland preservation	USA	CE	18	1,688	Acres	NA
Lindhjem	2007	Forests	Scandinavia	CV, CE	50	72	Hectares	Insensitive
Ojea et al.	2010	Forests	Worldwide	TC, RC, CV, others	65	172	ln(ha)	Negative
Ojea and Loureiro	2011	Ecosystems and diversity	USA, UK, some worldwide	CV	100	355	Absolute and relative changes in measurements in study	Insensitive (0.5) and Positive (0.5)
Richardson and Loomis[a]	2009	Threatened and endangered species	Worldwide	CV	31	67	% change	Positive
Smith and Osborne	1996	Visibility in national parks	USA	CV	5	115	% change in visibility	Positive

Table 3 (continued)

Study	Year	Commodity	Region	Method	Studies	Obs.	Size Measure	Scope
Woodward and Wui	2001	Wetlands	Worldwide	NFI, TC, RC, CV	39	65	ln(acres)	Negative
Zandersen and Tol	2009	Forests	Europe	TC	26	189	Hectares	Insensitive
Use								
Brander et al.[b]	2007	Coral reef ecosystems	Worldwide	TC, PF, NFI, CV	33	73	ln(km squared)	Positive
Johnston et al.	2005	Aquatic resources	USA	CV, others	34	81	Change in water quality	Positive
Van Houtven et al.	2007	Water quality	USA	CV, CE, TC, others	90	1,014	Water quality index	Positive
Poe et al.[c]	2001	Groundwater	USA	CV	14	208	% change in supply of water	Positive

Notes:
a. Richardson and Loomis (2009) is an update of Loomis and White (1996).
b. Brander et al. (2007) (coral reef ecosystems) is categorized as use because the studies examined WTP for entrance fees, not non-use values.
c. Poe et al. (2001) is an update of Boyle et al. (1994).

Key: CE choice experiment; CR contingent ranking; CV contingent valuation; HP hedonic prices; MP market prices; NFI net factor income; OC opportunity cost; PF production function; RC replacement cost; TC travel cost.

benefits, as these are entirely use in nature.[29] We have divided the updated Ojea and Loureiro table into meta-studies of largely or entirely non-use amenities and mixed use/non-use studies in one group and meta-studies of primarily or entirely use amenities in a second group.[30] For meta-studies of non-use and mixed non-use/use amenities, 5.5 studies find positive scope sensitivity and 6.5 find no or negative scope sensitivity.[31] The very mixed results of these scope meta-studies certainly do not support the view in the literature that most CV studies find scope sensitivity.

With respect to the primarily use category, the four meta-studies that focus largely or entirely on use amenities are Brander et al. (2007), which reviews studies estimating WTP for entry fees for coral reef ecosystems; Johnston et al. (2005), which reviews studies of water bodies that provide recreational benefits; Van Houten et al. (2007), which also reviews studies of the recreational benefits of water bodies; and Poe et al. (2001), which reviews studies of drinking water. All four of the meta-studies of amenities that are primarily use in nature find positive scope responsiveness. It is not surprising that studies of primarily use amenities find scope more frequently than studies of primarily non-use amenities, as respondents are much more likely to have well-formed utility functions for use amenities, particularly for amenities with which they have some experience purchasing or incurring costs (such as travel costs) to consume.

There are a number of flaws in the meta-studies that render their results unreliable for the purpose of evaluating whether CV studies pass scope. First, all the meta-studies in Table 3 include multiple data points from the same studies, including in some instances multiple data points from responses by the same respondent. Multiple WTPs elicited from the same respondent are "internal" WTPs.[32] As observations of this type are likely

[29] Meta-studies of recreational benefits include Walsh et al. (1984, 1992). Smith and Kaoru (1990), Sturtevant et al. (1998), Rosenberger and Loomis (2000a), Markowski et al. (2001), Bateman et al. (2003), Shrestha and Loomis (2003), and Van Houtven et al. (2007).

[30] We have not attempted to disentangle non-use from use WTP in studies of amenities that have both non-use and use values (estimates of WTP in such cases are commonly referred to as total use WTP).

[31] Ojea and Loureiro (2011) classify Johnston and Duke (2009) as finding negative scope sensitivity. However, while the coefficient of the scope variable is negative, the dependent variable is WTP/acre, so the sensitivity is negative only if total WTP declines with scope. The article does not provide data to determine if this is the case, so we change the classification of this study from negative sensitivity to N/A. Ojea and Loureiro conclude that Brouwer et al. (1999) find scope responsiveness. However, the relative size variable in the Brouwer et al. (1999) study is statistically not significant (p. 54) and the authors do not report whether the coefficient was positive or negative, so we report this study as scope insensitive.

[32] Some of the multiple estimates of WTP may be from split samples, and thus not subject to this criticism that they are internal.

to exhibit at least some scope sensitivity, including these contaminated observations in the meta-analysis will bias the overall results towards a finding of scope sensitivity.[33] A number of the studies included in the meta-scope reviews use choice experiments (CEs) as the elicitation methodology. Including observations from CE studies biases the results towards finding scope. WTPs inferred from these studies are based on choices of respondents who answer a series of questions with multiple choices. These responses are inherently "internal" in nature.

Second, the meta-studies conflate studies using a variety of non-CV methodologies for measuring WTP in addition to CV, including CE (as discussed above), the travel cost method, replacement cost, and hedonic analysis, among others. The only methodologies that are designed to infer WTP for non-use values are CV and CE – the others are either inappropriate or irrelevant for estimating non-use values. Only a handful of meta-studies are focused exclusively on CV.

Third, the studies do not correct for different cost scales across different studies. Cost scale has been found to have a strong positive correlation with WTP – see Cameron and Huppert (1989), Duffield and Patterson (1991), Cooper and Loomis (1992), Ryan and Wordsworth (2000), Hanley et al. (2005), Carlsson and Martinsson (2008), Mørkbak (2010), Kragt (2013), Prelec et al. (forthcoming) and Burrows, Dixon and Chan, Chapter 1 in this volume. For example, Carlsson and Martinsson (2008) find that an approximate doubling of the baseline cost range increase estimated WTP by two to three times; Burrows et al. find that increasing the cost scale by a factor of four increases estimated WTP by about three. As studies of "large" goods will tend to use cost scales that are larger than for studies of "small goods," not taking into account cost scale will bias a meta-study towards finding scope.

To illustrate the effect of cost scale on estimated WTPs, we examined the studies Ojea and Loureiro (2011) reviewed in their meta-analysis. Their analysis of scope effects is restricted to studies using area (measured in hectares) as the scope variable. We were unable to obtain the underlying database from the authors. We therefore reviewed the 109 studies they cited to identify all studies reporting hectares (or other measures convertible into hectares) as size variables. At least 70 of the 109 studies were not focused on environmental amenities measurable by area. For each of the studies involving benefits measured in areas, we collected information on

[33] Ojea and Loureiro (2011) is the only meta-scope study that recognizes this problem. The authors include a binary variable for split-sample observations vs within-subject observations. However, they do not interact this variable with the scope variable, so their approach does not test (and correct) for whether split-sample observations are more likely to exhibit scope than within-subject observations.

the area measurement (converted to hectares), reported WTPs and, where available, the cost scale. The reported WTPs were converted to 2016 US dollars. Many studies either did not report the cost scale or did not have a cost scale because they used an elicitation technique other than dichotomous choice. In addition, many of the studies did not provide area measurements and were focused on the number of land areas (such as forests, parks, and wetlands), as opposed to size.

We identified seven usable studies that reported areas, WTPs, and cost scale. One study (Petrolia and Kim, 2009) reports results for two hectare amounts and two separate cost scales for each.[34] A second study (McFadden, 1994) reports WTPs for 526,091 hectares, compared to size ranges of 100 hectares to 20,000 hectares for the other studies. McFadden (1994) also reported 26 WTP estimates using different elicitation and estimation methodologies. If authors reported multiple WTPs for the same amenity, we used the WTP that the authors indicated was preferred. As McFadden (1994) and Mill et al. (2007) do not indicate a preference, we used the average WTP reported in those studies. Table 4 summarizes the data for each study, ranked in order of size of affected amenity being valued. The range of hectares is huge, ranging from 100 hectares to 526,091 hectares. The range in cost scales is also huge, ranging from \$1.04–12.45 to \$3.65–3,653.89 (for the largest amenity measured in hectares). The range in WTPs is correspondingly large. In general, hectares, cost scale, and WTP seem to be broadly correlated, although the presence of an extremely large outlier would tend to drive the results of any econometric analysis, as is the case here.

Without the underlying data set and with only seven observations, we could not replicate the Ojea and Loureiro (2011) model including cost scale as an additional variable. Instead, we regressed WTP on hectares and the upper end of the cost scale, with and without a constant term and with and without the high hectare outlier (see Table 5). If the outlier is included and cost scale is excluded, hectares is highly significant, with or without a constant term. These equations viewed in isolation would suggest a highly significant scope effect. Cost scale is also highly significant if it is included without hectares, with or without a constant term. If both cost scale and hectares are included, the cost scale is significant and the hectares coefficient is insignificant in all of the estimated equations, with or without a constant term.

These results are suggestive, although there is a limit to how much can be inferred from seven studies. In these regressions, the cost scale seems to

[34] We also relied upon the average of the WTP estimates using the Turnbull, RE probit, income-bound RE probit, and high-bound RE probit methods for the two hectare amounts, per the authors' methodology on p. 144.

Table 4 Key metrics from CV articles with total cost scale (2016 $US) and amenity measured in area (hectares)

Authors	Year	Environmental Amenity	Hectares	Cost Scale Low ($)	Cost Scale High ($)	WTP ($)
Mill et al.[a]	2007	Forest preservation	100	15.79	236.85	215.23
Hammitt et al.	2001	Wetlands preservation	153	3.10	434.19	106.75
Kwak et al.	2003	Forest preservation	871	1.04	12.45	1.62
Petrolia and Kim	2009	Barrier island restoration	946	85.40	424.77	196.58
Petrolia and Kim	2009	Barrier island restoration	2,415	216.26	1,083.54	258.69
Loomis et al.	1994	Forest and animal preservation	2,833	3.22	483.36	145.15
Kniivila et al.	2002	Forest reservation	20,000	10.54	189.98	60.98
McFadden[b]	1994	Wilderness area preservation	526,091	3.65	3,653.89	1,217.50

Notes:

a. The WTP estimate in Mill et al. (2007) is an average of the means for the personal mixed, natural, and pine forests.

b. The WTP estimate in McFadden (1994) is an average of the 26 WTP estimates he gives for the 526,091-hectare Selway-Bitterroot Wilderness. WTP estimates range between $28.46 and $3,256.23.

Table 5 Regressions examining the relationship between WTP, hectares, and cost scale high

Variables	(1) WTP	(2) WTP	(3) WTP	(4) WTP	(5) WTP	(6) WTP	(7) WTP	(8) WTP	(9) WTP	(10) WTP	(11) WTP	(12) WTP
Hectares	0.0021**	0.002***	−0.0039	0.0060					0.0007	0.0003	0.0021	0.0003
	(0.00)	(0.00)	(0.01)	(0.01)					(0.00)	(0.00)	(0.00)	(0.00)
Cost Scale High					0.3249***	0.3274***	0.2102**	0.2856**	0.2202*	0.2842**	0.2016**	0.2843**
					(0.02)	(0.02)	(0.06)	(0.04)	(0.06)	(0.04)	(0.07)	(0.05)
Constant	127.6858**		150.6140**		5.7740		49.2112		42.5905		60.7100	
	(35.24)		(39.32)		(27.78)		(33.07)		(32.52)		(39.89)	
Observations	8	8	7	7	8	8	7	7	8	8	7	7
R-squared	0.9520	0.9011	0.1066	0.0853	0.9775	0.9853	0.6833	0.8811	0.9857	0.9876	0.7110	0.8813
RMSE	92.53	153	89.69	162.4	63.40	58.90	53.40	58.56	55.31	58.51	57.03	64.10

Notes:
Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Key (column no.):

1. WTP regressed on hectares with both the outlier and a constant in the model.
2. WTP regressed on hectares with outlier, but no constant in the model.
3. WTP regressed on hectares without the outlier, but with the constant in the model.
4. WTP regressed on hectares with no outlier and no constant in the model.
5. WTP regressed on CostScaleHigh with both the outlier and a constant in the model.
6. WTP regressed on CostScaleHigh with outlier, but no constant in the model.
7. WTP regressed on CostScaleHigh without the outlier, but with a constant in the model.
8. WTP regressed on CostScaleHigh with no outlier and no constant in the model.
9. WTP regressed on hectares and CostScaleHigh with both the outlier and a constant in the model.
10. WTP regressed on hectares and CostScaleHigh with outlier, but no constant in the model.
11. WTP regressed on hectares and CostScaleHigh without the outlier, but with the constant in the model.
12. WTP regressed on hectares and CostScaleHigh with no outlier and no constant in the model.

affect the WTP estimates in all versions of the estimations, while the size variable (hectares) is only significant when the cost scale or the cost scale and the outlier are omitted. These results support the hypothesis that cost scale variations could be an important explanatory variable in the analysis of scope effects in meta-studies. At a minimum, the effect of cost scales needs to be taken into account in future meta-studies.

SUMMARY OF THE STATE OF THE SCOPE LITERATURE

Based mostly on the 1997 Carson article and the (inconclusive) results of meta-studies, the currently prevailing conventional wisdom in the environmental literature is that most CV studies exhibit scope sensitivity, although the literature is largely silent with respect to the issue of the adequacy of scope (aside from arguing that low scope elasticity can be explained by sharply declining marginal utility). Heberlein et al. (2005) state that "[t]he scope test. . .is a fairly sure way of enhancing the credibility of one's study, since most CV studies pass scope tests."[35] Carson and Hanemann (2005) state that: "[t]he empirical evidence is that there is some sensitivity to scope for a wide range of goods." Kling et al. (2012), citing the 1997 Carson survey paper and three meta-studies,[36] conclude that "scope effects are typically present in well-executed studies."[37] Haab et al. (2013) cite the 2012 DMT paper and several meta-studies[38] and conclude that "CVM studies do, in fact, tend to pass a scope test."[39]

Whitehead (2016) goes further and concludes that "CVM studies that pass the scope test produce results that are most useful for policy analysis. CVM studies that do not pass the scope test should be critically examined for behavioral anomalies. . .before the CVM is determined to be a valuation method that cannot measure preferences."[40] In other words, the author implies that the burden of proof with respect to the use of CVM falls on those questioning CVM, as the accumulated evidence supports the conclusion that CVM studies "tend to pass the scope test."[41] As the brief

[35] Heberlein et al. (2005), p. 20. Ironically, for the four different amenities that are analyzed in the Heberlein study, most of the internal scope tests reported fail and one of the four external scope tests fails.
[36] Smith and Osborne (1996), Brouwer et al. (1999), and Ojea and Loureiro (2011).
[37] Kling et al. (2012), p. 19.
[38] Smith and Osborne (1996), Richardson and Loomis (2009), and Ojea and Loureiro (2011).
[39] Haab et al. (2013), p. 608.
[40] Whitehead (2016), p. 21.
[41] Ibid., p. 18.

history above has shown, the accumulated evidence consists only of the 1997 Carson survey (which was incomplete, error-prone, and now dated), the 2012 DMT paper (which includes internal scope tests and hardly supports the conclusion that most external CV tests pass scope), and a potpourri of meta-studies with mixed results and that use designs that are biased towards finding scope (e.g., by not accounting for the effects of cost scale and by including "internal" WTPs).

REVIEW OF SCOPE RESULTS REPORTED IN CV STUDIES

In this chapter, we focus on the frequency with which CV estimates of WTP for non-use environmental goods and services demonstrate sensitivity to scope and on whether scope effects that are found are adequate (or plausible). We have compiled a comprehensive and up-to-date survey of all CV studies of environmental goods and services that contain an external scope test. Our analysis of these studies focuses on those that include estimates of WTP for non-use environmental goods and services. As many environmental goods have elements of both use value and non-use value (for example, some respondents may have a positive WTP for a clean lake because the knowledge that the lake is clean provides non-use utility, while other respondents may derive use utility from a clean lake because they like to swim in it, and yet other respondents may derive both use value and non-use value), we also include CV studies that derive WTP for goods and services that have both significant non-use utility as well as use utility (classification of utilities as use or non-use obviously requires some subjective judgment).[42] In addition to the searches of the literature, we also cross-referenced citations in all the articles we identified and included studies that also tested for scope.

For reasons discussed earlier, we include in our survey only studies that perform external scope tests. Carson's 1997 paper includes a table that nicely highlights the interaction between external and internal scope tests (reproduced in Table 6). This table shows three goods, A, B, and C, whose values are nested (A > B > C). The valuation sequences are denoted by I, II, and III. The order in which the goods are offered is indicated by

[42] To assemble our database of studies related to testing scope in CV surveys, we searched the EVRI (Environmental Valuation Reference Inventory) and NOEP (National Ocean Economics Program – Middlebury College) databases, government websites and publication sources (including NOAA, EPA, and the US Fisheries and Wildlife Agency, among others), academic websites (including Richard Carson's invaluable website for collected studies RePEc.org, the mammoth bibliography in Carson, 2012, EBSCO, Econlit, and Google Scholar).

Table 6 Reproduction of Carson (1997) Table I

Subsample I	Subsample II	Subsample III
A_I^1		
B_I^2	B_{II}^1	
C_I^3	C_{II}^2	C_{III}^1

Note: Goods A, B, and C are nested, with good A being the "largest" good.

a subscript 1, 2, or 3. It is assumed that all of the goods are normal and that all are substitutes of each other. The above taxonomy lends itself to a number of sequencing and scope tests. However, the only tests that are purely external are those along the diagonal: $A_I^1 \geq B_{II}^1 \geq C_{III}^1$. The other tests ($A_I^1 \geq B_I^2 \geq C_I^3$ and $B_{II}^2 \geq C_{II}^3$) are internal tests in that in one or more of the estimates respondents are being asked to state a value of a good after already having given a value of another good that is senior or junior to the good in question in a normal valuation sequence.

We focus exclusively on split-sample tests along the diagonal. By analyzing only responses to the first choice question posed to respondents, we minimize the influence of context and anchoring effects. The other tests in Carson's taxonomy would be valid tests of scope if human beings were robots impervious to suggestion and ignored signals from the survey, but in the real world any answers by human respondents are affected by what they have already seen. Context and anchoring effects of prior questions (or cues in the current question) undermine the usefulness of survey responses. Sequencing effects have been widely reported in the literature in which WTP for an item is higher if it is placed first in a list than later: see, among many examples, Randall et al. (1981), Boyce et al. (1989), Boyle et al. (1990), Samples and Hollyer (1990), Boyle et al. (1993), Hoehn and Loomis (1993), Halvorsen (1996), DuPont (2003), and Bateman et al. (2004, 2006). DuPont (2003) reports that her "review of the literature on question order indicates that its effect may be more strongly felt in cases where passive use WTP values are being sought as opposed to active use values."[43]

Our focus is consistent with the NOAA panel's conclusion that "We must reject one possible approach, that of asking each respondent to express willingness to pay to avert incidents of varying sizes; the danger is that embedding will be forcibly avoided, still without realism."[44] As Bateman et al. (2004) also observe, "it is widely recognized that passage

[43] Dupont (2003), p. 325.
[44] Arrow et al. (1993), p. 27.

of internal tests is relatively facile and possibly related to the observation that respondents may simply be trying to be 'internally consistent' in their reported values."[45]

We identified 111 studies published between 1983 and 2016 that include external scope tests of environmental goods; 104 of these studies present scope tests of non-use or mixed non-use/use environmental goods and services. We excluded a small number of studies that asked for WTP for reducing the probability of health effects from environmental and other risks, as it has been reported that human beings have difficulty with assessing small probabilities; these studies are also use oriented, which was not the focus of our analysis. Table 7 presents summary information on the 111 studies in our database, including information on the environmental amenity, whether the goods surveyed are use, non-use, or non-use mixed with some use, and our conclusions about whether each study passes or fails a scope test or has mixed results. The classifications of studies into use, non-use, and mixed non-use/use were based on subjective judgment. Most, but not all, papers report test statistics that allow significance tests on the pass/fail results. We use $p < 0.05$ as the threshold for significance; for cases in which no test statistics are provided we accept the conclusions of the authors about significance. Of course, a finding in a study of statistically significant scope responsiveness is not enough to establish validity – there must also be an assessment of whether the size of the scope test is adequate or plausible. If scope is low it is possible that there is a large element of warm glow in the measured WTP for the amenity.

Several studies report both internal and external scope tests. For studies that report both external and internal scope tests, we focus only on the results of the external scope tests: for example, if the external scope test fails, but the internal scope tests passes, we report the study as failing an external scope test (as passing an internal scope test may simply be the result of respondents attempting to appear internally consistent). A number of studies have a perverse finding that an internal test fails but an external test passes. For example, in the Giraud et al. (1999) study of Mexican spotted owls the internal scope test fails and the external scope test passes. Heberlein et al. (2005) report that the majority of his internal scope tests fail, but spearfishing passes an external scope test. Day and Mourato (1998) report that the pilot survey in their study fails an internal scope test but that the full field survey passes an external scope test. We report all these studies as passing scope. This is a very conservative method, as failing an internal scope test is either a symptom of a failure of the survey design or an indication that respondents are inconsistent in their answers.

[45] Bateman et al. (2004), p. 83.

Table 7 Scope test summary

Author	Date	Result Pass/Fail (Mixed: % Pass)	Commodity	Use
Ahearn et al.	2006	M (25%)	Protecting grassland birds in the Central Plains region	Non-use
Ahlheim et al.	2014	F	Plant species and rainforest in Yunnan, China	Non-use
Alvarez-Farzio et al.*	1999	P	Preserving two environmentally sensitive areas in Scotland	Both
Araña and León*	2008	M (33%)	Rehabilitating walking paths in Gran Canaria, Spain	Both
Banzhaf et al.	2011	P	Southern Appalachian Mountains ecosystem services	Non-use
Banzhaf et al.	2006	F	Acid rain: quality of water and fish populations, bird species, and tree species (600 lakes versus 900 lakes)	Both
Bateman et al.	2004	M (83%)	Open access lake improvements located in grounds of University of East Anglia	Both
Bennett et al.*	1998	P	Dryland salinity in Upper South East region of South Australia	Non-use
Bergstrom and Stoll	1987	F	Farmland protection in Greenville County, South Carolina	Both
Berrens et al.	1996	P	Protecting minimum instream flows in New Mexico: silvery minnow versus 11 threatened species	Non-use
Berrens et al.	2000	P	Protecting minimum instream flows in New Mexico: silvery minnow versus 11 threatened species	Non-use
Binger et al.	1995	P	Preservation of the Selway-Bitterroot Wilderness and 57 federal wilderness areas	Non-use
Bliem and Getzner	2008	F	Ecological restoration in Danube river basin	Non-use
Bliem and Getzner	2012	F	River restoration along the Austrian Danube	Non-use
Blomquist and Whitehead	1998	P	Preserving wetland areas in Kentucky	Non-use
Bowker and Didychuk	1994	P	Preserve units of Moncton area farmland	Non-use
Boxall et al.	2012	M (83%)	Marine mammal species in St. Lawrence Estuary	Non-use

Boyle et al.	1994	F	Preventing waterfowl deaths in the Central Flyway	Non-use
Brookshire et al.	1983	F	Benefit of hunting big horn sheep or grizzly bears in five or 15 years	Use
Brown and Duffield	1995	M (50%)	Protect instream flow in either one or five Montana rivers	Non-use
Brown et al.	1995	F	Preservation of natural areas in Fort Collins	Both
Carson	1997	F	Preventing birds from being killed	Non-use
Carson and Mitchell*	1993	P	National versus regional water quality	Both
Carson et al.	1994a	P	Preservation of Australia's Kakadu conservation zone from mining activity	Non-use
Carson et al.	1994b	P	Program to accelerate natural restoration of injured resources due to PCB and DDT contamination in Southern California Bight	Both
Carson et al.*	1990	P	Air quality in Cincinnati	Both
Carson and Mitchell	1995	P	Preventing injuries from large open pit mine in Kakadu Conservation Zone	Non-use
Caudill et al.	2011	F	Saginaw bay in Michigan	Non-use
Chapman et al.	2009	F	Alum treatments to prevent algae growth in the Illinois River system and Tenkiller Lake	Both
Choe et al.	1996	M (33%)	Improving water quality of rivers and sea in Davao, Philippines	Both
Christie	2001	M (17%)	Recreation opportunities, Grampian region, Scotland	Use
Day and Mourato	1998	P	Maintaining river water quality in Beijing	Both
Desvousges et al.	1992	F	Preventing waterfowl deaths in the Central Flyway	Non-use
Desvousges et al.	1993	F	Preventing waterfowl deaths in the Central Flyway	Non-use
Desvousges et al.	2016	F	Alum treatments to prevent algae growth in Oklahoma	Both
Desvousges et al.	2012	F	River and lake water clarity in Oklahoma	Both
Diamond et al.	1993	F	Preservation Selway, Washakie, and Bob Marshall wilderness areas	Non-use

Table 7 (continued)

Author	Date	Result Pass/Fail (Mixed: % Pass)	Commodity	Use
DuPont	2003	M (17%)	Improvements to three recreational activities (swimming, fishing, and boating in Hamilton Harbor, Ontario, Canada)	Use
DuPont	2013	F	Reclaimed wastewater in Canada	Use
Eom and Larson	2006	P	Water quality improvement in the Man Kyoung River in South Korea	Both
Fischhoff et al.*	1993	F	River clean-up in Pittsburgh area	Both
Fredman*	1995	F	Protecting the white-backed woodpecker in Sweden	Non-use
Gerrans*	1994	F	Preservation of the Jandakot wetlands	Non-use
Gilbert et al.*	1991	F	Lye Brook Wilderness Area in Southwestern Vermont	Both
Gillespie and Bennett	2011	F	Marine Protected Areas in South-West Marine Region, Australia	Non-use
Giraud et al.	1999	P	Mexican spotted owl and 62 regional threatened and endangered species	Non-use
Gong and Baron	2011	P	Four endangered species, four health risks	Both
Goodman et al.*	1998	F	Non-use value of natural coastal environment in England	Non-use
Hanemann*	2005	P	Saving at-risk birds (the bird study revisited)	Non-use
Heberlein et al.	2005	P	(1) Water quality, (2) Wisconsin's wild wolf population, (3) Indian spearfishing, (4) biodiversity	Both
Hicks et al.*	2004	M (47%)	Oyster reef restoration in Chesapeake Bay (asked to general public and fishermen)	Both
Hite et al.	2002	P	Subsidizing variable-rate technology to reduce polluted river runoff for the Mississippi River Basin (in Mississippi)	Non-use
Hoevenagel	1996	P	Greenhouse effect, depletion of ozone layer, deforestation, acid rain, surface water pollution, animal manure problem	Both

	Hsee and Rottenstreich	2004	M (50%)	Saving pandas	Non-use
	Huang et al.	1997	M (50%)	Quality improvement of Pamlico and Albemarle sounds (recreation areas) in North Carolina	Both
	Jakobsson and Dragun*	2001	P	Conserving endangered species and the Leadbeater's possum in Victoria, Australia	Non-use
	Jakus*	1992	P	Gypsy moth protection	Both
	Jin et al.	2010	M (25%)	Marine turtle conservation in multiple Asian cities	Non-use
	Kahneman*	1986	M (50%)	Clean-up to preserve fishing in Muskoka, Haliburton, and all Ontario	Use
	Kahneman and Knetsch	1992	M (50%)	Preserving wilderness, protecting wildlife, providing parks, preparing for disasters, controlling air pollution, insuring water quality, routine treatment of industrial wastes; 12 embedding pairs	Both
	Kahneman and Ritov*	1994	F	Headline method – multiple	Both
	Krieger*	1994	P	Anglers' WTP for changes in Michigan's public health advisory	Use
	Longo et al.	2012	F	Climate change mitigation	Non-use
	Loomis and Ekstrand	1997	F	Preserving Mexican spotted owl and 62 endangered and threatened species	Non-use
	Loomis et al.	1993	M (50%)	Protections of forests in Southeast Australia	Non-use
	Macmillan et al.	1996	F	Recovery/damage scenarios from reduced acid rain deposition	Non-use
	Macmillan and Duff*	1998	P	Native woodland restoration in the UK	Non-use
	Magnussen	1992	M (50%)	Reduce Norwegian nutrient leaching to the North Sea	Both
	McDaniels et al.	2003	P	Benefits of fisheries enhancement on rivers in British Columbia	Both
	McFadden	1994	F	Preservation of the Selway-Bitterroot Wilderness in Northern Idaho	Non-use

Table 7 (continued)

Author	Date	Result Pass/Fail (Mixed: % Pass)	Commodity	Use
McFadden and Leonard	1993	F	Preservation of the Selway-Bitterroot Wilderness	Non-use
Metcalfe	2012	P	Improvement of water quality in England and Wales	Use
Mullarkey and Bishop	1999	P	Wisconsin wetlands	Non-use
Mullarkey	1997	P	Wisconsin wetlands	Non-use
Navrud	1989	F	Reduced sulfur depositions to protect freshwater fish populations	Both
Nunes and Schokkaert	2003	M (50%)	Protection of wilderness and recreation areas in Alentejo Natural Park, Portugal	Both
Pattison et al.	2011	M (50%)	Wetland retention and restoration in Manitoba	Non-use
Poe et al.	2005	M (43%)	Mexican spotted owl and 62 regional threatened and endangered species	Non-use
Pouta	2005	F	Forest regeneration cutting policy in Finland	Both
Pouta*	2003	F	Preservation of nature conservation areas and environmentally oriented forest management in Finland	Both
Powe and Bateman	2004	P	Protect wetlands (nested area versus total) in Broadland, Eastern England	Non-use
Rahmatian	1986	F	Visibility in Grand Canyon National Park	Both
Rathnayake*	2016	P	Ecotourism in Kawdulla National Park, Sri Lanka	Use
Ready et al.	1997	P	Preventing a decrease in the number of horse farms in Kentucky	Non-use
Ressurreição et al.	2011	M (50%)	Marine taxa in Azores archipelago	Non-use
Rollins and Lyke	1998	F	Creating parks in Canada's Northwest Territories	Non-use
Rowe et al.	1992	P	Oil spill clean-up and prevention programs	Both

Samples and Hollyer*	1990	F	Preserving humpback whales and Hawaiian monk seals in Hawaii	Non-use
Schkade and Payne	1994	M (50%)	Preserving migratory waterfowl in the Central Flyway of the USA (the bird study)	Non-use
Schulze et al.	1995	F	Upper Clark Fork River, Montana, River Basin Restoration	Both
Smith et al.	1997	P	Controlling marine debris on beaches and coastal areas in New Jersey and North Carolina	Both
Smith et al.	2005	F	Improvements of Eastern regional haze	Both
Stanley	2005	P	Riverside fairy shrimp versus all local endangered species	Non-use
Stevens et al.	1997	P	Movie passes and restoration of Atlantic salmon	Both
Stevens et al.*	1995	P	Preserving wetlands in New England	Non-use
Streever et al.	1998	F	Wetland conservation in New South Wales, Australia	Non-use
Svedsäter	2000	F	Rain forests in South America, endangered wild animals, air pollution in central London, global warming	Both
Tanguay et al.*	1993	F	Maintaining woodland caribou population in Northwestern Saskatchewan	Non-use
Total Value Team	2015, 2015b	P	*Deepwater Horizon recovery*	Non-use
Veisten et al.	2004a	M (71%)	Endangered species preservation in Norwegian forests	Non-use
Veisten et al.	2004b	F	Norwegian forests endangered species preservation	Non-use
Vo and Huynh	2014	F	Groundwater protection program	Both
Welsh et al.*	1994	P	Glen Canyon Dam downstream recreation, hydropower, and passive-use values	Use
Welsh et al.*	1995	M (30%)	Preserving resources of Colorado River downstream of Glen Canyon Dam	Both

Table 7 (continued)

Author	Date	Result Pass/Fail (Mixed: % Pass)	Commodity	Use
White et al.*	1997	F	Conservation of two otter species	Non-use
Whitehead and Cherry*	2007	M (67%)	Green energy program in North Carolina (yielding improved air quality in western North Carolina mountains)	Both
Whitehead and Finney*	2003	F	Preserving submerged marine cultural resources (historic shipwrecks)	Non-use
Whitehead et al.	1998	P	Quality improvement of Pamlico and Albemarle sounds (recreation areas) in North Carolina	Both
Whitehead et al.	2007	F	Purchase and managing additional acres of coastal marshes in Saginaw Bay	Non-use
Whitehead et al.*	2009	F	Purchase and managing additional acres of coastal marshes in Saginaw Bay	Non-use
Wilson	2000	P	Wolves, Chippewa Indian spearfishing, biological diversity, and water quality in Wisconsin	Both

Note: * = excludes relevant test statistics.

To accommodate the different adaptations and interpretations of the same survey by various authors, we also track the occurrence of multiple authors using the same dataset. There are 11 instances in which one underlying survey provides the basis for two papers (i.e., both papers used the same underlying data) and one instance in which one survey has been the basis of three papers. To avoid over-weighting of the underlying survey, in our tabulations we assign partial weights to studies based on identical data sets (0.5 each for cases in which two papers rely on the same underlying survey and 0.33 each for the case in which three papers rely on the same underlying survey). These surveys include the following:

- Boyle et al. (1994) present survey results on WTP for preventing waterfowl deaths in the Central Flyway that fail scope. Hanemann (2005) administers a survey that is nearly identical to the Boyle et al. survey and finds scope sensitivity.
- Schkade and Payne (1994) also present survey results on WTP for preventing waterfowl deaths in the Central Flyway that fails scope in the 2,000 to 20,000 increment, and appears to pass (although no test statistics are explicitly provided) in the 20,000 to 200,000 increment. We consider this paper to be half-pass, half-fail. Carson (1997) shows that removal of certain data outliers results in a directional passage in both increments, but a statistical failure in both as well. This paper is a fail, weighted as 0.5 with the Schkade and Payne (1994) result.
- Giraud et al. (1999) on the Mexican spotted owl and 62 regional endangered and threatened species report a scope test as passing; using the same data and different statistical procedures, Poe et al. (2005) report mixed scope effects.
- Wilson (2000) presents survey results and external scope tests on WTP for wolves (N/A) and water quality (pass) in Wisconsin.[46] Using the same data, Heberlein et al. (2005) report external scope tests for wolves and for water quality that pass. As we state earlier, we treat the results for wolves as N/As because we cannot rule out that respondents' true underlying marginal WTP for more wolves might in fact be zero (or negative).
- Whitehead et al. (2007, 2009) and Caudill et al. (2011) all use the same survey data set with respect to protection of marshland and report that the scope tests fail.

[46] Wilson also reports internal scope tests for Chippewa spearfishing and biological diversity.

CV studies differ widely across many dimensions, including the data that are included in the analysis, the elicitation technique (such as dichotomous choice [DC], double-bounded [DB], payment card [PC], open ended [OE], and bidding), the payment vehicle (such as taxes, higher prices, and voluntary contributions), features of the survey designs and the survey design itself, and the analytical models and statistical procedures used to estimate WTPs. It is not feasible to deconstruct each of the studies to determine the preferred method in each context. If the authors report multiple scope tests and indicate which methods and results are preferred, in our tabulations we report the scope results that are favored by the authors; if the authors do not indicate which they feel are preferred, we report all the external scope tests and include them in our tabulations with fractional allocation.[47] This is the case for the factors described below.

VARIATIONS IN THE DATA INCLUDED IN THE ANALYSIS

One of the shortcomings of the CV literature is that different authors vary with respect to what data are included or excluded in their analyses. Some authors include all the data in a survey, although this is more the exception than the rule. It is common for authors to exclude so-called "protest" votes – some exclude all zeros, and some exclude zero answers if the respondents provide certain answers to debriefing questions. In some cases (but much less commonly than for protest votes), authors exclude WTP answers that are too high to be credible (for example, Rowe et al., 1992 exclude WTPs that are over 1% of a respondent's income and Smith et al., 2005 exclude the highest 5% of WTPs) or answers by respondents

[47] In one case, we report the conclusions of the authors even though we disagree with their methodology. Whitehead et al. (1998, water quality improvement in Albermarle [A] and Pamlico [P] sounds in North Carolina), report split-sample estimates of WTPs for A only and for A + P that pass scope; however, both samples included both A and P residents; P residents valuing A may have WTPs for A that are low or zero, which would lead to the result that WTP (A) is less than WTP (A + P) even if the respondents are not sensitive to scope. We report this study as passing scope. Loomis et al. (2009, reductions in acres burned by wildfires in California, Florida, and Montana) report three split-sample estimates for residents in each state. The authors report a logistic regression model in which acreage is significant. However, the acreage variable is fixed in size for each state, so this variable is equivalent to a dummy variable for each state. If California residents place a higher value than Montana residents and Montana residents place a higher value than Florida residents on reducing wildfires for reasons unrelated to acreages involved, the study would have found the same effects. We tabulate this study as N/A, as it does not appear to include a true external scope test.

that are judged to be yea-sayers as a result of follow-up questions. Authors sometimes exclude answers that the respondents indicate are uncertain. MacMillan and Duff (1998) include only respondents who say they are willing to pay taxes. Some studies interview local and non-local respondents separately. There is no accepted practice with respect to what data to include and what to include. As WTP estimates can be greatly influenced by what data are included or excluded, this gives analysts enormous latitude to vary which data to include (subsequently contributing to a wide variation in their possible results). Some examples of studies with idiosyncratic approaches to the data include the following:

- Powe and Bateman (2004) report a CV study of wetlands that fails scope, but after removing respondents who, based on debriefing questions, do not consider the program realistic, find scope. We report this study as passing scope in our tabulations.
- Nunes and Schokkaert (2003), in a study of wilderness and recreation areas in national parks in Portugal, report results that fail the scope test in their base case but that pass the scope test after removal of respondents who they conclude from econometric analysis of attitudinal questions gain utility from warm glow. We report this study as half-pass, half-fail.
- Christie (2001) reports scope tests using the full survey data set and 5% truncated samples (removing high bid outliers). As the author infers that the truncated sample is preferred, we include the results from only the truncated sample in our summary results.
- Welsh et al. (1995) report results for a national sample and a local sample for alternative flows for Glen Canyon. They conclude scope passes for the national sample, but the local marketing area results generally fail scope. The authors conclude that "combined with the lack of demonstrated sensitivity to scope for the marketing area in the pilot survey, the case for stating that the marketing area has passed advanced tests is somewhat weaker than for the national sample." We report mixed results and allocate passes and fails fractionally.

VARIATIONS IN SURVEY DESIGN

Survey designs also vary widely, and some surveys have unusual or one-off designs that generate results that are different from "standard" surveys. Examples include the following:

- Hanemann (2005) conducts a survey with a similar design as the Desvousges et al. (1993) bird study, with the only difference being that the amenity is prevention of bird deaths expressed in percentages instead of absolute amounts, and reports finding scope sensitivity. We record this study as passing scope, and include this study in our tabulations with a 50% weight (so that this result offsets the Desvousges et al. finding of scope failure, which also has a 50% weighting factor).
- Huang et al. (1997) report single-bounded dichotomous choice results that fail scope and double-bounded dichotomous choice results that pass scope. We weigh each result with a 50% weight.[48]
- Fischhoff et al.'s (1993) evidence for scope insensitivity with direct estimates and scope sensitivity when respondents make paired comparisons.[49] We report this study as failing scope.
- Veisten et al. (2004a, 2004b) report results from models based on both OE and PC approaches, without indicating which is preferred. These all fail scope.

VARIATIONS IN ANALYTICAL MODELS AND STATISTICAL PROCEDURES

The analytical models and statistical procedures used to derive WTPs vary significantly across studies, and include both parametric and non-parametric techniques. There is no accepted practice for which techniques to use and no science for which are preferable other than measuring which provide the best statistical fits to the data. Some studies use parametric models that include covariates and others do not. Some studies correct for heteroscedasticity and others do not. Some studies use a Box-Cox transformation to correct for skewed bids. The parametric functions used differ across studies and within studies. For cases in which authors report multiple models and statistical procedures, in our tabulation of scope conclusions we accept the versions that the authors favor; in cases in which authors present alternative versions without indicating which are

[48] This paper is weighted as 0.5 with the other Pamlico and Albemarle text (Whitehead et al., 1998), so overall it yields 0.25 pass, 0.25 fail in our fractional tabulations.

[49] Fischhoff et al. state that "The effects of task simplification are studied by contrasting the performance of subjects using two response modes: 1) *direct estimates*, i.e., assigning dollar values to individual goods; and 2) *paired comparisons*, i.e., choosing the better of two competing goods. . . Our results revealed a very high degree of embedding with the direct estimates, but only a moderate degree with the paired comparisons" (p. 212; original emphasis).

preferred, we accept all the results reported, with fractional weighting for our tabulations. Examples include the following:

- Poe et al. (2005) present results based on a linear logit model and also a non-parametric truncated Kriström model, some of which pass and some of which fail. We report these results fractionally.
- Choe (1996) presents alternative results for two scope scenarios based on an OLS model (pass at $p < 0.05$), a hazard/Weibull model (pass at $p < 0.10$), and a probit model (insignificant). We report these results as one-third pass, two-thirds fail.

INCONSISTENT STATISTICAL SIGNIFICANCE RESULTS

A number of studies report scope variables that are significant or tests that indicate that the coefficients of the estimating equations are significantly different for different scopes, even though the reported confidence intervals for the WTPs overlap. We report results such as these as passes even if the reported WTPs fail scope.[50] Examples include the following:

- Ready et al. (1997) report a logistic regression in which scope dummies are significant, but some of the confidence intervals for the reported WTPs overlap. We report this study as passing scope.
- Bowker and Didychuk (1994) report that the acreage variable is significant in an OLS regression of factors explaining WTP, but report mixed scope results with respect to WTPs for different acreage amounts. We record this study as passing scope.
- Smith et al. (1997) report that the 95% confidence intervals for WTPs for different levels of beach clean-up overlap, but that "using a likelihood ratio test to test whether the parameters of the hazard functions describing respondents' stated choices for each of the four default beach scenes were equal, we reject the null hypothesis with a

[50] The one exception is the case of Banzhaf et al. (2006). This paper includes a statistically significant and positive coefficient for the scope variable presented in Table 4; however, the 90% confidence intervals overlap using the Weibull model for their preferred symmetric all-econometric-controls option, indicating a fail, but pass at the 10% level using the log-normal model for the same option (the 95% test statistics, our threshold, are not provided). Last, the authors present their "cautious, best defensible estimates of the mean WTP" after adjusting for discount factors at $48–107 per year, per household in the base scenario, and $54–159 for the scope scenario. We report this paper as a fail.

p-value = 0.05."[51] We record this study as passing scope, even though the reported WTPs do not pass scope.

How Frequently do CV Studies Pass a Scope Test?

For the 111 studies of environmental goods in our study, 41 (37%) report scope tests that pass scope tests, 46 (41%) report scope tests that fail, and 24 (22%) report mixed results (see Table 7). After fractional allocation of the mixed tests to the pass and fail categories, 52.3 (47%) of the scope tests pass and 58.7 (53%) fail.[52] After appropriately weighting the studies that are based on common underlying data, 45.1 (46%) pass and 52.9 (54%) fail.

Table 7 also reports scope pass/fail results for environmental non-use and mixed use/non-use goods. Focusing on the latter category, after correction for weighting factors and fractional allocations, 40.3 (45%) of the studies pass scope and 48.7 (55%) fail scope. Excluding studies that do not report statistical tests to allow determination of significance of the scope tests, 31 (47%) pass and 35.2 (53%) fail the scope tests.

If the ability to detect scope effects in CV studies is a function of the quality of the survey, one would expect to find a positive relationship between survey quality and the percentage of studies passing a scope test. One way to test this is to determine if the percentage of CV studies that pass scope has increased over time, as presumably the quality of CV surveys has increased over time (see Figure 1). Over the 15-year period 1987–2001, 51% of the studies reported scope tests that pass (based on a fractional allocation) and 49% reported scope tests that failed. For the 15-year period from 2002 to the present, the pass percentage declined to 41% and the fail percentage increased to 59%. As applications of CVM have become increasingly sophisticated over time, with most studies following the guidelines developed over the history of CVM for meeting acceptable standards of quality, the percentage of reported scope tests that pass scope has actually dropped sharply over the past 15 years relative to the prior 15-year period. Presumably, the sharp decline in the pass rate is not the result of a decline in the quality of the studies reported, nor can it be the result of deteriorating human cognitive facilities.

Another indicator of quality of a study might be whether it is published in a peer-reviewed academic journal. The results for published studies are

[51] Smith et al. (1997), p. 239.

[52] If a study reports mixed results, we weight each result reported fractionally, with the weights adding up to one. For example, if a study reports three tests that pass and two that fail, in our tabulations the study is reported as 0.6 pass and 0.4 fail.

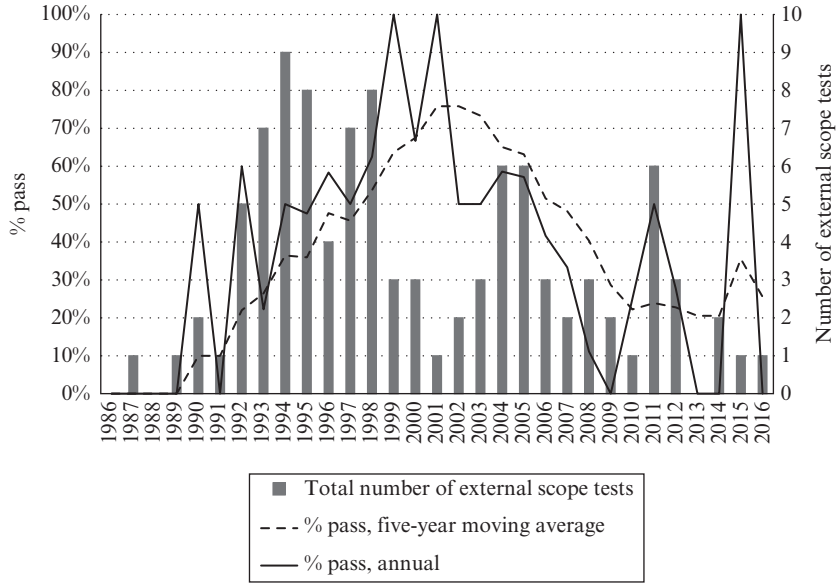

Figure 1 Passes as a percentage of total in environmental external scope tests, five-year moving average

similar to the results for all of the studies. On a raw count basis for studies published in peer-reviewed literature, 37% have scope tests that pass, 18% have mixed scope tests, and 46% have scope tests that fail. After fractional allocation, the results are pass (48%) and fail (52%), similar to the results for all studies.

DO SCOPE TESTS DEMONSTRATE ADEQUATE SENSITIVITY TO SCOPE?

The NOAA Panel's concern over scope focused in part on whether CV studies demonstrate "adequate" sensitivity to scope. There is no analytical basis for determining how much scope sensitivity is "adequate" or "plausible." Defining sensitivity in terms of elasticity of scope, a scope elasticity of somewhat less than one would be viewed by most economists as plausible (in the presence of declining marginal utility, most environmental goods and services would have scope elasticities of less than one). It does not seem plausible that utility for environmental goods would be

easily satiated, as might be the case for many use goods (such as ice cream cones at the beach to give an extreme example). If the value of preserving one species is $10, we think most observers would expect the value of preserving a second similar species would also be close to $10 (and similarly for acres of clean beaches and numbers of pristine lakes or forests). A judgment of plausibility of a scope test will understandably differ for different observers.

In the case of many of the CV surveys that pass a scope test, as in the case of the Araña and León (2008) study of walking path kilometers, WTP for the larger good is only slightly larger than WTP for the smaller good. For example, in Poe et al. (2005) the authors ask 369 survey respondents whether they would pay various costs to protect the Mexican spotted owl. A separate group of 363 respondents are asked about their WTP to protect the Mexican spotted owl and 62 additional threatened and endangered species. The study finds that mean WTP for protection of the Mexican spotted owl is $99.93 and mean WTP for protection of the Mexican spotted owl and 62 additional threatened and endangered species is $130.42. This study passes a directional scope test, but is this estimated scope "adequate"? The results imply that, on average, the typical respondent values each of the 62 additional species by an amount that is less than 1% of their valuation of the Mexican spotted owl.

There is a tendency for CV studies to demonstrate a kink in the demand for an environmental amenity at the base level of that amenity that is defined in the survey (see our earlier discussion of diminishing marginal utility and bounded substitution). For example, a study by Rollins and Lyke (1998) of national parks in Canada's Northwest Territories informs respondents that 29 of 39 natural regions are already national parks, and then asks respondents about WTP for additional parks. The reported WTPs are $105.45 for one additional park, $161.85 for two more parks, $191.07 for four more parks, and $188.44 for ten more parks. The increment from one to two additional parks is significant, but the increments from two to four and from four to ten additional parks are not. If these WTPs are correct representations of respondent utility, the demand curve for parks rapidly becomes flat in the range of 29–31 parks. It is conceivable (but not credible, in our view) that this is accurate for this particular study, but is it credible that in most such studies the kink in the demand curve is at or just above the base quantity in the study?

The majority of CV studies have scope tests that are either "categorical" in nature or do not provide information on the change in quantity in the survey, and therefore do not lend themselves to a quantitative analysis. Table 8 presents the non-failing (pass and mixed) scope tests in 24 "quantitative" scope studies that do lend themselves to an analysis of quantitative scope

effects.[53] Table 9 reports scope results only for those tests that pass scope. For each scope test we identify the good, the baseline quantity, the scope quantity, the baseline WTP, the scope WTP, the scope elasticity, and the weight of the scope test used in our weighted average calculations. The reported WTPs are means, as this is the metric most commonly reported in CV studies and is available for virtually all the studies in our survey.

Figure 2 presents a histogram of the scope elasticity effects for the median result in each study using all reported scope results, including the fractional passes from papers that report mixed effects. For the 21 results we could include in this analysis,[54] nine have scope elasticities of less than 0.10 and 12 have scope elasticities of less than 0.2; only three have scope elasticities above 0.5. It is worth bearing in mind that only 46% of reported scope tests pass; if, say, scope elasticities above 0.5 are plausible, this means that only about 7% (3/21 × 46%) of reported scope studies pass with "adequate" scope. If an elasticity of 0.2 is the threshold for plausibility, only about 20% (9/21 × 46%) of reported scope tests pass with adequate scope.

The frequency of limited scope elasticities documented in this study suggests that warm glow is an important element of measured WTP for environmental amenities. Diminishing marginal utility and bounded substitution cannot credibly explain the low levels of scope responsiveness across all of the CV studies we measured. This justification would mean that these factors dominate WTP valuations at levels just larger than the smaller environmental good presented in each individual paper, chosen by each individual author, across a range of unique goods. Put differently, how is it possible that in our review of the entire scope literature we found so few examples of WTPs for incremental goods that are even commensurate with the initial valuation? Traditional scope tests (i.e., non-adding-up tests) cannot determine whether findings of limited or no scope sensitivity are explained by "warm glow," diminishing marginal utility, part-whole bias, disagreement with the survey's implied probability of provision of the larger good, cognitive shortcomings, or any other explanations that have been proffered.

[53] Two studies – Giraud et al. (1999) and Poe et al. (2005) – use the same survey data of Mexican spotted owls and each find elasticity of WTP of less than 0.2. These are treated as one study in Figure 2. We exclude two studies with absurdly high scope elasticities that are not credible: Ready et al. (1997, horse farms) with a median scope elasticity of 65.98 and Pattison et al. (2011, wetlands) with a scope elasticity of 3.8.

[54] Giraud et al. (1999) and Poe et al. (2005) are each weighted as 0.5.

Table 8 Quantitative tests of scope elasticities

Author	Year	Commodity	Base Quantity	Scope Quantity	Base WTP ($)	Scope WTP ($)	Elasticity
McDaniels et al.	2003	Rivers	1	10	36	95	0.18
Berrens et al.	2000	Miles of river	170	1,000	24	52	0.24
Berrens et al.	2000	Miles of river	170	1,000	26	58	0.24
Berrens et al.	2000	Miles of river	170	1,000	24	52	0.24
Berrens et al.	2000	Miles of river	170	1,000	662	79,328	24.33
Berrens et al.	2000	Miles of river	170	1,000	206	1,819	1.6
Berrens et al.	2000	Miles of river	170	1,000	26	72	0.35
Carson et al.	1994	Years of recovery period	50	15	63	34	0.66
Bowker and Didychuk	1994	Acres of farmland	71,250	95,000	78	86	0.29
Bowker and Didychuk	1994	Acres of farmland	47,500	71,250	68	78	0.32
Bowker and Didychuk	1994	Acres of farmland	23,750	47,500	49	68	0.38
Binger et al.	1995	Wilderness areas	1	57	29	79	0.03
Giraud et al.	1999	Specie	1	62	48	118	0.02
Araña and León	2008	Kilometers	30	300	20	28	0.05
Araña and León	2008	Kilometers	100	300	26	28	0.05
Araña and León	2008	Kilometers	30	100	20	26	0.13
Poe et al.	2005	Specie	1	62	100	130	0.01
Poe et al.	2005	Specie	1	62	116	156	0.01
Eom and Larson	2006	Reductions in biochemical oxygen demand level	4	7	10	18	1
Macmillan and Duff	1998	Forests to be saved	1	2	53	67	0.26
Macmillan and Duff	1998	Forests to be saved	1	2	35	67	0.91

Macmillan and Duff	1998	Forests to be saved	1	2	69	109	0.58
Macmillan and Duff	1998	Forests to be saved	1	2	51	109	1.14
Stanley	2005	Specie	1	32	30	56	0.03
Stanley	2005	Specie	1	32	31	64	0.04
Stanley	2005	Specie	1	32	31	64	0.03
Stanley	2005	Specie	1	32	21	38	0.03
Stanley	2005	Specie	1	32	28	59	0.04
Stanley	2005	Specie	1	32	25	52	0.04
Hsee and Rottenstreich	2004	Pandas	1	4	11	22	0.31
Nunes and Schokkaert	2003	Areas	1	2	3,500	6,000	0.71
Nunes and Schokkaert	2003	Areas	1	2	2,600	6,000	1.31
Hanemann	2005	Percentage of at-risk bird population saved	100,000	1,000,000	22	34	0.06
Jakobsson and Dragun	2001	Number of species	1	700	29	267	0.01
Hicks et al.	2004	Acres	1,000	10,000	11	24	0.13
Hicks et al.	2004	Acres	5,000	10,000	22	24	0.12
Hicks et al.	2004	Acres	2,500	5,000	20	22	0.09
Hicks et al.	2004	Acres	1,000	2,500	11	20	0.53
Ahearn et al.	2006	Bird species increase percentage	10	50	12	14	0.04
Ahearn et al.	2006	Bird species increase percentage	10	50	11	13	0.06
Ahearn et al.	2006	Bird species increase percentage	10	50	12	14	0.03
Ahearn et al.	2006	Bird species increase percentage	10	50	11	12	0.03

Table 8 (continued)

Author	Year	Commodity	Base Quantity	Scope Quantity	Base WTP ($)	Scope WTP ($)	Elasticity
Hite et al.	2002	Percentage decreases in river runoff	0	0	47	50	0.06
Alvarez-Farzio et al.	1999	Environmentally sensitive area	1	2	13	36	1.68
Alvarez-Farzio et al.	1999	Environmentally sensitive area	1	2	25	36	0.43
Brown and Duffield	1995	Rivers	1	5	7	12	0.21
Brown and Duffield	1995	Rivers	1	5	10	18	0.19
Whitehead et al.	1998	Bodies of water	1	2	113	137	0.21
Berrens et al.	1996	Miles of river	170	1,000	29	90	0.43
Schkade and Payne	1994	Bird deaths prevented	2,000	200,000	63	122	0.11
Schkade and Payne	1994	Bird deaths prevented	2,000	200,000	69	134	0.10
Diamond et al.	1993	Wilderness areas	1	57	29	79	0.03
Loomis et al.	1993	Hectares of land	6,000	70,000	57	103	0.08
Loomis et al.	1993	Hectares of land	70,000	122,000	103	100	−0.04

Table 9 Quantitative tests of scope elasticities (pure passes only)

Author	Year	Commodity	Base Quantity	Scope Quantity	Base WTP ($)	Scope WTP ($)	Elasticity
McDaniels et al.	2003	Rivers	1	10	36	95	0.18
Berrens et al.	2000	Miles of river	170	1,000	24	52	0.24
Berrens et al.	2000	Miles of river	170	1,000	2	58	0.24
Berrens et al.	2000	Miles of river	170	1,000	24	52	0.24
Berrens et al.	2000	Miles of river	170	1,000	662	79,328	24.33
Berrens et al.	2000	Miles of river	170	1,000	206	1,819	1.6
Berrens et al.	2000	Miles of river	170	1,000	26	72	0.35
Carson et al.	1994	Years of recovery period	50	15	63	34	0.66
Bowker and Didychuk	1994	Acres of farmland	71,250	95,000	78	86	0.29
Bowker and Didychuk	1994	Acres of farmland	47,500	71,250	68	78	0.32
Bowker and Didychuk	1994	Acres of farmland	23,750	47,500	49	68	0.38
Binger et al.	1995	Wilderness areas	1	57	29	79	0.03
Giraud et al.	1999	Specie	1	62	48	118	0.02
Eom and Larson	2006	Reductions in biochemical oxygen demand level	4	7	10	18	1
Macmillan and Duff	1998	Forests to be saved	1	2	53	67	0.26
Macmillan and Duff	1998	Forests to be saved	1	2	35	67	0.91
Macmillan and Duff	1998	Forests to be saved	1	2	69	109	0.58
Macmillan and Duff	1998	Forests to be saved	1	2	51	109	1.14
Stanley	2005	Specie	1	32	30	56	0.03
Stanley	2005	Specie	1	32	31	64	0.04

Table 9 (continued)

Author	Year	Commodity	Base Quantity	Scope Quantity	Base WTP ($)	Scope WTP ($)	Elasticity
Stanley	2005	Specie	1	32	31	64	0.03
Stanley	2005	Specie	1	32	21	38	0.03
Stanley	2005	Specie	1	32	28	59	0.04
Stanley	2005	Specie	1	32	25	52	0.04
Hanemann	2005	Percentage of at-risk bird population saved	100,000	1,000,000	22	34	0.06
Jakobsson and Dragun	2001	Number of species	1	700	29	267	0.01
Hite et al.	2002	Percentage decreases in river runoff	0	0	47	50	0.06
Alvarez-Farzio et al.	1999	Environmentally sensitive area	1	2	13	36	1.68
Alvarez-Farzio et al.	1999	Environmentally sensitive area	1	2	25	36	0.43
Whitehead et al.	1998	Bodies of water	1	2	113	137	0.21
Berrens et al.	1996	Miles of river	170	1,000	29	90	0.43

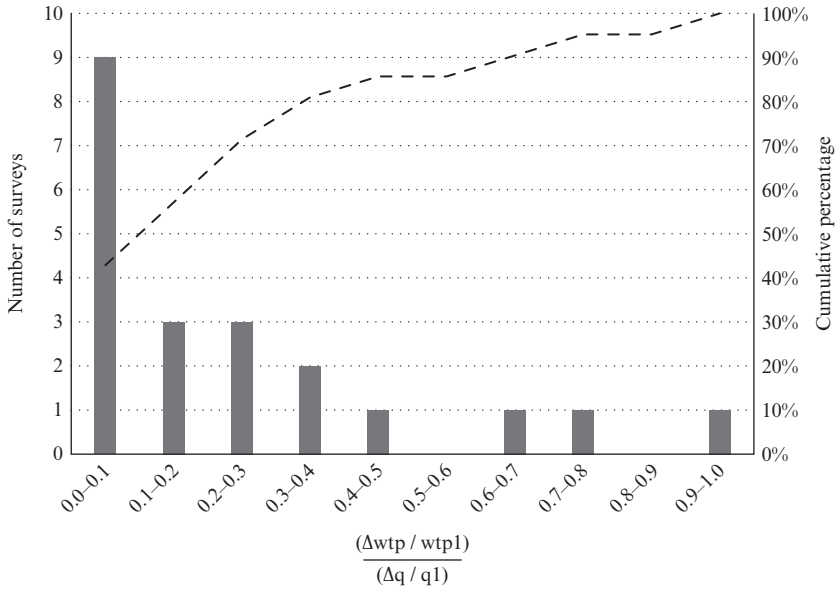

Figure 2 Elasticity of WTP median result per paper

ARE SCOPE PASS AND FAIL RESULTS AFFECTED BY MEASURABLE CHARACTERISTICS OF THE SURVEYS?

As we have shown, it is not uncommon for CV studies to fail a scope test. It would be useful to know how CV studies could be designed to pass scope (and presumably more reliably reflect underlying utility). In an effort to shed some light on this issue, we collected information on a variety of characteristics of the CV studies we reviewed. These include: sample size, year of publication (year of survey was not reported for many studies), elicitation methodology (dichotomous choice, multiple-bounded dichotomous choice, open ended, and payment card), survey method (in person, telephone, Internet, and mail), frequency of payment (single payment, annual, other), public/private, presence of cheap talk, presence of a budget reminder, certainty correction, removal of outliers, presence of dissonance minimization, and whether WTP was reported for a gain or a loss. The overall quality of the survey might be an important factor, but we have no direct measure of quality. Certain proxies are available, however. CV studies have presumably improved over time, so year of publication might be correlated with survey quality. Larger sample sizes and response rates

may also be indicators of quality. We do have data on sample size, but most CV studies that report scope tests do not provide data on response rates for the scope split samples.

Table 10 presents the simple correlations between each individual variable and a dummy variable for pass. All of the simple correlation coefficients are small. The largest positive effects are for the presence of certainty corrections, WTP measured as a gain, and year of survey. In all of these cases, the direction of the effect is in the expected direction. Surprisingly, sample size has a negative simple correlation with pass. "Public" is negatively correlated (as expected if respondents have a less well-defined utility for public goods vs private goods). "Mandatory" payment and "open-ended" survey mode are negatively correlated with pass, neither of which is expected.

We also estimated probit multivariate regressions, reported in Table 10. With all explanatory variables included (column 1), the R^2 is only 0.07. The only significant variables are "public" (negative coefficient), and "online" survey methodology (negative); the interpretation of "public" is suspect because only four studies in our database estimated WTPs for private goods. Column 2 reports the results after dropping variables for which there were either few observations (certainty corrections, public, removal of outliers) or for which we were uncertain about the reporting accuracy of the variables (e.g., "budget reminder" and "cheap talk," as many articles were silent on these aspects of the survey). The only variable that is significant in this version is "online" survey mode. Column 3 reports the results in which the "public" variable is dropped (as all but four studies are public). In addition, the "mandatory" and "online" payment vehicle and "mail" and "phone" survey mode variables are dropped. In this version, only "online" survey mode is significant. Column 4 shows the results in which payment frequency "once" and "periodic" are dropped. In this version, only the "online" survey mode variable is significant. It is notable that in no version is the time trend or sample size significant, so the two variables that might be proxies for survey quality seem to have no explanatory power.

In sum, very little of the variance of pass/fail is explained by the measurable characteristics that we could identify for the CV studies we reviewed. The only variable that is consistently significant is "online" survey mode. This provides weak evidence that administrating a survey online may have some slight effect on the ability of the survey to pass scope.

Table 10 Regressions examining factors that affect scope insensitivity

Variable	Correlation	(1) Pass	(2) Pass	(3) Pass	(4) Pass
Intercept		−27.6331	−36.3914	−27.5609	−21.0349
		(45.8366)	(43.9923)	(42.0119)	(40.6707)
Samplesizetotal	−0.022825866	−0.00003	0.000024	0.000017	0.000017
		(0.000153)	(0.000146)	(0.00014)	(0.00014)
Gain	0.186205405	0.397	0.3144	0.3199	0.287
		(0.2914)	(0.2798)	(0.2707)	(0.267)
Pub_year	0.17989491	0.0141	0.0185	0.0134	0.0104
		(0.0229)	(0.022)	(0.021)	(0.0203)
Em_multiple	0.072486809	0.4179	0.6248	0.4187	0.4096
		(0.3967)	(0.3764)	(0.3633)	(0.3534)
Em_open	−0.265319614	−0.6479	−0.5093	−0.5017	−0.4565
		(0.3994)	(0.3807)	(0.3551)	(0.3471)
Em_ paymentcard	0.138246794	0.0353	0.3429	0.1588	0.1962
		(0.4569)	(0.4022)	(0.3883)	(0.3836)
Em_single	0.112646765	Omitted	Omitted	Omitted	Omitted
Sm_person	0.101937002	−0.3016	−0.426	−0.0826	−0.1001
		(0.4268)	(0.4129)	(0.3037)	(0.3008)
Pv_voluntary	0.091034085	0.1519	0.1868	0.0571	−0.0168
		(0.3515)	(0.3399)	(0.3197)	(0.308)
Pv_NA	0.148859071	0.1656	0.3295	0.1664	0.1999
		(0.5519)	(0.5461)	(0.5291)	(0.5264)
Pf_once	0.013647689	0.9123	0.8841	0.393	
		(0.8718)	(0.8668)	(0.823)	
Pf_periodic	−0.012771008	0.9749	1.0128	0.5673	
		(0.8583)	(0.8516)	(0.8121)	
Pf_NA	−0.001102309	Omitted	Omitted	Omitted	
Public	−0.100940019	−1.5108*	−1.5162		
		(0.8155)	(0.8072)		
Pv_indirect	0.066137319	0.1577	0.303		
		(0.484)	(0.4479)		
Pv_mandatory	−0.206242136	Omitted	Omitted		
Sm_online	0.041647614	−1.733**	−1.4775**	−1.0901**	−1.0513*
		(0.6965)	(0.6312)	(−0.5543)	(0.5543)
Sm_mail	−0.045961431	−0.4985	−0.5421		
		(0.3952)	(0.3783)		
Sm_phone	−0.058656592	Omitted	Omitted		
Budget	0.018154951	−0.4105			
		(0.3379)			
Cheap	0.076138778	1.0716			
		(0.9813)			
Certain	0.189879845	0.5652			
		(1.0945)			

Table 10 (continued)

Variable	Correlation	(1) Pass	(2) Pass	(3) Pass	(4) Pass
Outlier	0.016710393	0.7365* (0.3864)			
R^2		0.0718	0.0526	0.0338	0.0317
Max-rescaled R^2		0.1907	0.1397	0.0897	0.0843
AIC		168.78	167.41	167.779	164.459
Observations		324	324	324	324
Sum of weights		111	111	111	111

Note: Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

CONCLUSIONS

A fundamental tenet of consumer utility theory is that for most goods utility increases as consumption increases. The 1993 NOAA Panel report concluded that the findings of a CV study would be "unreliable" if the survey exhibited "inadequate responsiveness to the scope of the environmental insult." Four of the Panel members later stated that "had the panel thought that something as straightforward as statistical measurability were the proper way to define sensitivity, then we would (or should have) opted for language to that effect. A better word than 'adequate' would have been 'plausible'."

In spite of the importance of demonstrating that CV studies are adequately responsive to scope, and thus CV estimates of WTP are accepted as exhibiting rationality, there has been very little systematic review in the literature of the extent to which CV studies pass scope and the extent to which the scope findings in those studies that pass are "adequate" or "plausible." The literature also fails to tell us how CV studies can be designed to elicit rational WTPs that exhibit adequate scope responsiveness. An early study by Carson (1997) implies that 31 studies that he identifies all passed scope; his survey does not include at least 35 additional studies, most of which fail scope, and the studies he identifies include four that do not have a scope test, three that have mixed results, and two that fail scope. A later study by Desvousges et al. (2012) reports that more studies have mixed results or fail scope than pass scope.

A number of meta-studies review scope tests across studies; for amenities that are non-use or mixed non-use/mix, more report negative scope findings than positive scope findings. In addition, these studies are flawed

because they include multiple data points from the same respondents and do not correct for the effects of the cost scale. In addition, most of the meta-studies include studies that use other techniques than CV, the only methodology that is appropriate for estimating WTP for non-use amenities.

We find that approximately 54% of a comprehensive set of 111 environmental CV studies fail to demonstrate scope. Studies dated in the last 15 years (2002–16) have a higher failure rate (59%) than studies dated in the prior 15-year period (1987–2001), suggesting that the significant advances in the sophistication of CV methodology from its early days have not improved the ability of this methodology to estimate rationally sound WTPs.

For the minority of papers that do pass, few of the studies that pass scope exhibit scope elasticities in a range that, in our judgment, is plausible. Just under half of the studies that provide scope results that can be quantified exhibit scope elasticities of less than 0.2, and less than one-fourth of the studies exhibit scope elasticities over 0.5.

The only characteristic of the studies we reviewed that seems to be weakly associated with passing scope is online administration of the survey. There is a need for additional analysis to determine how to improve the CV methodology so that it exhibits results that plausibly pass scope, a basic criterion of rationality.

REFERENCES

Ahearn, M., K. Boyle, and D. Hellerstein (2006), "Designing a contingent valuation study to estimate the benefits of the conservation reserve program," in A. Alberini and J. Kahn (eds), *Handbook of Contingent Valuation*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Ahlheim, M., O. Frör, G. Langenberger, and S. Pelz (2014), "Chinese urbanites and the preservation of rare species in remote parts of the country: The example of Eaglewood," *Environmental Economics*, **5**(4), 32–43.

Alvarez-Farizo, B., N. Hanley, R.E. Wright, and D. Macmillan (1999), "Estimating the benefits of agri-environmental policy: Econometric issues in open-ended contingent valuation studies," *Journal of Environmental Planning and Management*, **42**(1), 23–43.

Amiran, E. and D.A. Hagen (2010), "The scope trials: Variation in sensitivity to scope and WTP with directionality bounded utility functions," *Journal of Environmental Economics and Management*, **59**(3), 293–301.

Andreoni, J. (1990), "Impure altruism and donations to public goods: A theory of warm glow giving," *Economic Journal*, **100**(401), 464–77.

Araña, J.E. and C.J. León (2008), "Do emotions matter? Coherent preferences under anchoring and emotional effects," *Ecological Economics*, **66**(4), 700–711.

Arrow, K., E.E. Leamer, H. Schuman, and R. Solow (1994), *Comments on Proposed NOAA/DOI Regulations on Natural Resource Damage Assessment*, US Environmental Protection Agency.

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schumanj (1993), *Report of the NOAA Panel on Contingent Valuation*, December 1, 2016 at http://www.economia.unimib.it/DATA/moduli/7_6067/materiale/noaa%20report.pdf.

Banzhaf, H.S., D. Burtraw, D. Evans, and A. Krupnick (2006), "Valuation of natural resource improvements in the Adirondacks," *Land Economics*, **82**(3), 445–64.

Banzhaf, H.S., D. Burtraw, S. Chung, D.A. Evans, A. Krupnik, and J. Siikamaki (2011), "Valuation of ecosystem services in the Southern Appalachian Mountains," paper presented at the Annual Meeting of the Association of Environmental and Resource Economists.

Barrio, M. and M.L. Loureiro (2010), "A meta-analysis of contingent valuation forest studies," *Ecological Economics*, **69**(5), 1020–30.

Bateman, I. (2003), "Contrasting conventional with multi-level modeling approaches to meta-analysis: Expectation consistency in U.K. woodland recreation values," *Land Economics*, **79**(2), 235–58.

Bateman, I.J., A.A. Lovett, and J.S. Brainard (2003), *Applied Environmental Economics. A GIS Approach to Cost–Benefit Analysis*, Cambridge, UK: The Press Syndicate of the University of Cambridge.

Bateman, I.J., M. Cole, S. Georgiou, and D. Hadley (2006), "Comparing contingent valuation and contingent ranking: A case study considering the benefits of urban river water quality improvements," *Journal of Environmental Management*, **79**(3), 221–31.

Bateman, I.J., M. Cole, P. Cooper, S. Georgiou, D. Hadley, and G.L. Poe (2004), "On visible choice sets and scope sensitivity," *Journal of Environmental Economics and Management*, **47**(1), 71–93.

Bennett, J., M. Morrison, and R. Blamey (1998), "Testing the validity of responses to contingent valuation questioning," *The Australian Journal of Agricultural and Resource Economics*, **42**(2), 131–48.

Bergstrom, J.C. and J.R. Stoll (1987), "A test of contingent market bid elicitation procedures for piecewise valuation," *Western Journal for Agricultural Economics*, **12**(2), 104–8.

Berrens, R.P., P. Ganderton, and C.L. Silva (1996), "Valuing the protection of minimum instream flows in New Mexico," *Journal of Agricultural and Resource Economics*, **21**(2), 294–309.

Berrens, R.P., A.K. Bohara, C.L. Silva, D. Brookshire, and M. McKee (2000), "Contingent values for New Mexico instream flows: With tests of scope, group-size reminder and temporal reliability," *Journal of Environmental Management*, **58**(1), 73–90.

Binger, B.R., R.F. Copple, and E. Hoffman (1995), "The use of contingent valuation methodology in natural resource damage assessments: Legal fact and economic fiction," *Northwestern University Law Review*, **89**(3), 1029–53.

Bishop, Richard, Kevin, Boyle, Michael Welsh, Robert Baumgartner and Pamela Rathburn (1986), Glen Canyon Dam Releases and Downstream Recreation: An Analysis of User Preferences and Economic Values. Final Report to the Bureau of Reclamation. HBRS Madison, Wisconsin.

Bliem, M. and M. Getzner (2008), *Valuation of Ecological Restoration Benefits*

in the Danube River Basin Using Stated Preference Methods – Report on the Austrian Case Study Results, Klagenfurt: Institute for Advanced Studies Carinthia, Department of Economics, Klagenfurt University.
Bliem, M. and M. Getzner (2012), "Willingness-to-pay for river restoration: Differences across time and scenarios," *Environmental Economics and Policy Studies*, **14**(3), 241–60.
Blomquist, G.C. and J.C. Whitehead (1998), "Resource quality information and validity of willingness to pay in contingent valuation," *Resource and Energy Economics*, **20**(2), 179–96.
Boman, M. and G. Bostedt (1999), "Valuing the wolf in Sweden: Are benefits contingent on the supply?," *Topics in Environmental Economics*, **17**, 157–74.
Bowker, J.M. and D.D. Didychuk (1994), "Estimation of the nonmarket benefits of agricultural land retention in Eastern Canada," *Agricultural and Resource Economics Review*, **23**(2), 218–25.
Boxall, P.C., W.L. Adamowicz, M. Olar, G.E. West, and G. Cantin (2012), "Analysis of the economic benefits associated with the recovery of threatened marine mammal species in the Canadian St. Lawrence Estuary," *Marine Policy*, **36**(1), 189–97.
Boyce, R.R., T.C. Brown, G.D. McClelland, G.L. Peterson, and W.D. Schulze (1989), "Experimental evidence of existence value in payment and compensation contexts," paper presented at the Joint Meeting of the Western Committee on the Benefits and Costs of Natural Resource Planning and the Western Regional Science Association, San Diego, CA.
Boyle, K.J., S.D. Reiling, and M.L. Phillips (1990), "Species substitution and question sequencing in contingent valuation surveys evaluating the hunting of several types of wildlife," *Leisure Sciences*, **12**(1), 103–18.
Boyle, K.J., M.P. Welsh, and R.C. Bishop (1993), "The role of question order and respondent experience in contingent valuation studies," *Journal of Environmental Economics and Management*, **25**(S), 8,064–99.
Boyle, K.J., W.H. Desvousges, F.R. Johnson, R.W. Dunford, and S.P. Hudson (1994), "An investigation of part-whole biases in contingent valuation studies," *Journal of Environmental Economics and Management*, **27**(1), 64–83.
Brander, L.M., P. van Beukering, and H.S.J. Cesar (2007), "The recreational value of coral reefs: A meta-analysis," *Ecological Economics*, **63**(1), 209–18.
Brookshire, D.S., L.S. Eubanks, and A. Randall (1983), "Estimating option prices and existence values for wildlife resources," *Land Economics*, **59**(1), 1–15.
Brouwer, R., I.H. Langford, I.J. Bateman, and R.K. Turner (1999), "A meta-analysis of wetland contingent valuation studies," *Regional Environmental Change*, **1**(1), 47–57.
Brown, T.C. and J.W. Duffield (1995), "Testing part-whole valuation effects in contingent valuation of instream flow protection," *Water Resources Research*, **31**(9), 2,341–51.
Brown, T.C., S.C. Barro, M.J. Manfredo, and G.L. Peterson (1995), "Does better information about the good avoid embedding effect?," *Journal of Environmental Management*, **44**(1), 1–10.
Cameron, T.A. and D.D. Huppert (1989), "OLS versus ML estimation of non-market resource values with payment card interval data," *Journal of Environmental Economics and Management*, **17**(3), 230–46.
Carlsson, F. and P. Martinsson (2008), "How much is too much? An investigation of the effect of the number of choice sets, context dependence and the choice

of bid vectors in choice experiments," *Environmental Resource Economics*, **40**, 165–76.
Carson, R.T. (1994), *Prospective Interim Lost Use Value Due to PCB and DDT Contamination in the Southern California Bight, Volume 1*, report by Natural Resource Damage Assessment Inc. and Industrial Economics Inc. to the NOAA.
Carson, R.T. (1997), "Contingent valuation and tests of insensitivity to scope," in R.J. Kopp, W. Pommerhene, and N. Schwartz (eds), *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*, Boston, MA: Kluwer.
Carson, R.T. (2012), "Contingent valuation: A practical alternative when prices aren't available," *Journal of Economic Perspectives*, **26**(4), 27–42.
Carson, R.T. and W.M. Hanemann (2005), "Contingent valuation," in K.G. Maler and J.R. Vincent (eds), *Handbook of Environmental Economics, Volume 2: Valuing Environmental Changes*, Amsterdam: North-Holland.
Carson, R.T. and R.C. Mitchell (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Washington, DC: Resources for the Future.
Carson, R.T. and R.C. Mitchell (1993), "The value of clean water: The public's willingness to pay for boatable, fishable, and swimmable quality water," *Water Resources Research*, **29**(7), 2,445–54.
Carson, R.T. and R.C. Mitchell (1995), "Sequencing and nesting in contingent valuation surveys," *Journal of Environmental Economics and Management*, **28**(2), 155–73.
Carson, R.T., R.C. Mitchell, and P.A. Ruud (1990), "Valuing air quality improvements: Simulating a hedonic equation in the context of a contingent valuation scenario," in C. Mathi (ed.), *Visibility and Fine Particles*, Pittsburgh, PA: Air and Waste Management Association.
Carson, R.T., L. Wilks, and D. Imber (1994a), "Valuing the preservation of Australia's Kakadu Conservation Zone," *Oxford Economic Papers*, **46**(Supplement), 727–49.
Carson, R.T., W.M. Hanemann, R.J. Kopp, J.A. Krosnick, R.C. Mitchell, S. Presser, P.A. Ruud, and V.K. Smith (1994b), *Prospective Interim Lost Use Value Due to PCB and DDT Contamination in the Southern California Bight*, report to National Oceanic and Atmospheric Administration by Natural Resource Damage Assessment Inc.
Caudill, S.B., P.A. Groothius, and J.C. Whitehead (2011), "The development and estimation of a latent choice multinomial logit model with application to contingent valuation," *American Journal of Agricultural Economics*, **93**(4).
Chapman, D.J., R.C. Bishop, W.M. Hanemann, B.J. Kanninen, J.A. Krosnick, E.R. Morey, and R. Tourangeau (2009), *Natural Resource Damages Associated with Aesthetic and Ecosystem Injuries to Oklahoma's Illinois River System and Tenkiller Lake. Expert Report for State of Oklahoma*, accessed December 1, 2016 at https://pprg.stanford.edu/wp-content/uploads/9-Natural-resource-damages-associated-with-aesthetic-and-ecosystem-injuries-to-Oklahomas-Illinois-river-system.pdf.
Choe, K., D. Whittington, and D.T. Lauria (1996), "The economic benefits of surface water quality improvements in developing countries: A case study of Davao, Philippines," *Land Economics*, **72**(4), 519–37.
Christie, M. (2001), "A comparison of alternative contingent valuation elicitation treatments for the evaluation of complex environmental policy," *Journal of Environmental Management*, **62**(3), 255–69.

Cooper, J. and J. Loomis (1992), "Sensitivity of willingness-to-pay estimates to bid design in dichotomous choice contingent valuation models," *Land Economics*, **68**(2), 211–24.

Cummings, R.G., D.S. Brookshire, and W.D. Schulze (1986), *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, Totowa, NJ: Rowman & Allanheld.

Day, B. and S. Mourato (1998), "Willingness to pay for water quality maintenance in Chinese rivers," *CSERGE Working Paper*, Centre for Social and Economic Research on the Global Environment, University College London and University of East Anglia.

Desvousges, W., K. Mathews, and K. Train (2012), "Adequate responsiveness to scope in contingent valuation," *Ecological Economics*, **84**(1), 121–8.

Desvousges, W., K. Mathews, and K. Train (2016), "An adding-up test on contingent valuations of river and lake quality," *Land Economics*, forthcoming.

Desvousges, W.H., F.R. Johnson, R.W. Dunford, S.P. Hudson, K.N. Wilson, and K.J. Boyle (1992), *Measuring Non-use Damages Using Contingent Valuation: An Experimental Evaluation of Accuracy*, Research Triangle Institute Monograph No. 92-1, Research Triangle Park, NC.

Desvousges, W.H., F.R. Johnson, R.W. Dunford, S.P. Hudson, K.N. Wilson, and K.J. Boyle (1993), "Measuring natural resource damages with contingent valuation: Tests of validity and reliability," in J.A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*, Amsterdam: Elsevier, pp. 91–164.

Diamond, P.A. (1996), "Testing the internal consistency of contingent valuation surveys," *Journal of Environmental Economics and Management*, **30**(3), 337–47.

Diamond, P.A., J.A. Hausman, G.K. Leonard, and M.A. Denning (1993), "Does contingent valuation measure preferences? Experimental evidence," in J.A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*, Amsterdam: Elsevier, pp. 41–89.

Duffield, J. and C. Neher (1991), *A Contingent Valuation Assessment of Montana Waterfowl Hunting*, report prepared for the Montana Department of Fish Wildlife, and Parks, Montana Bioeconomics, Missoula, MT.

Duffield, J.W. and D.A. Patterson (1991), "Field testing existence values: An instream flow trust fund for Montana rivers," paper presented at the Association of Environmental and Resource Economics during the Valuing Environmental Goods with Contingent Valuation session.

DuPont, D.P. (2003), "CVM embedding effects when there are active, potentially active and passive users of environmental goods," *Environmental and Resource Economics*, **25**(3), 319–41.

DuPont, D.P. (2013), "Water use restrictions or wastewater recycling? A Canadian willingness to pay study for reclaimed wastewater," *Water Resources and Economics*, **1**(1), 61–74.

Eom, Y.-S. and D.M. Larson (2006), "Improving environmental valuation estimates through consistent use of revealed and stated preference information," *Journal of Economics and Environmental Management*, **52**(1), 501–16.

Fischhoff, B., M.J. Quadrel, M. Kamlet, G. Loewenstein, R. Dawes, P. Fischbeck, S. Klepper, J. Leland, and P. Stroh (1993), "Embedding effects: Stimulus representation and response mode," *Journal of Risk and Uncertainty*, **6**, 211–34.

Fredman, P. (1995), "The existence of existence value – study of the economic benefits of an endangered species," *Journal of Forest Economics*, **1**(3), 307–27.

Gerrans, P. (1994), "An economic valuation of the Jandakot wetlands," *Occasional Paper No. 1*, Perth: Edith Cowan University.
Ghermandi, A., J.C.J.M. van der Bergh, L.M. Brander, H.L.F. de Groot, and P.A.L.D. Nunes (2008), "The economic value of wetland conservation and creation: A meta-analysis," *Nota di Lavoro No. 79.2008*, Fondazione Eni Enrico Mattei.
Gilbert, A., R. Glass, and T. More (1991), "Valuation of eastern wilderness: Extramarket measures of public support," in C. Paine, J. Bowker, and P. Reed (eds), *The Economic Value of Wilderness*, Ashville, NC: US Department of Agriculture, Forest Service.
Giraud, K.L., J.B. Loomis, and R.L. Johnson (1999), "Internal and external scope in willingness-to-pay estimates for threatened and endangered wildlife," *Journal of Environmental Management*, **56**(3), 221–9.
Gong, M. and J. Baron (2011), "The generality of the emotion effect on magnitude sensitivity," *Journal of Economic Psychology*, **32**(1), 17–24.
Goodman, S.L., W. Seabrooke, and S.A. Jaffry (1998), "Considering conservation value in economic appraisals of coastal resources," *Journal of Environmental Planning and Management*, **41**(3), 313–36.
Haab, T.C., M.G. Interis, D.R. Petrolia, and J.C. Whitehead (2013), "From hopeless to curious? Thoughts on Hausman's 'dubious to hopeless' critique of contingent valuation," *Applied Economic Perspectives and Policy*, **35**(4), 593–612.
Halvorsen, B. (1996), "Ordering effects in contingent valuation surveys: Willingness to pay for reduced health damage from air pollution," *Environmental and Resource Economics*, **8**(4), 485–99.
Hammit, J.K., L. Jin-Tan, and L. Jin-Long (2001), "Contingent valuation of a Taiwanese wetland," *Environment and Development Economics*, **6**(2), 259–68.
Hanemann, M. (2005), "The bird study revisited," presentation, UC Berkeley.
Hanley, N., R.E. Wright, and B. Alvarez-Farizo (2005), "Estimating the economic value of improvements in river ecology using choice experiments: An application to the Water Framework Directive," *Journal of Environmental Management*, **78**(2), 183–93.
Hausman, J. and W. Newey (2016), "Individual heterogeneity and average welfare," *Econometrica*, **84**(3), 1225–48.
Heberlein, T.A., M.A. Wilson, R.C. Bishop, and N.C. Schaeffer (2005), "Rethinking the scope test as a criterion for validity in contingent valuation," *Journal of Environmental Economics and Management*, **50**(1), 1–22.
Hicks, R.L., T.C. Haab, and D. Lipton (2004), *The Economic Benefits of Oyster Reef Restoration in the Chesapeake Bay*, report for the Chesapeake Bay Foundation.
Hite, D., D. Hudson, and W. Intarapapong (2002), "Willingness to pay for water quality improvements: The case of precision application technology," *Journal of Agricultural and Resource Economics*, **27**(2), 433–49.
Hjerpe, E., A. Hussain, and S. Phillips (2015), "Valuing type and scope of eco-system conservation: A meta-analysis," *Journal of Forest Economics*, **21**(1), 32–50.
Hoehn, J.P. and J.B. Loomis (1993), "Substitution effects in the valuation of multiple environmental programs," *Journal of Environmental Economics and Management*, **25**(1), 56–75.
Hoevenagel, R. (1996), "The validity of the contingent valuation method: Perfect and regular embedding," *Environmental and Resource Economics*, **7**(1), 57–78.

Hsee, C.K. and Y. Rottenstreich (2004), "Music, pandas, and muggers: On the affective psychology of value," *Journal of Experimental Psychology*, **133**(1), 23–30.

Huang, J.-C., T.C. Haab, and J.C. Whitehead (1997), "Willingness to pay for quality improvements: Should revealed and stated preference data be combined?" *Journal of Environmental Economics and Management*, **34**(3), 240–55.

Irwin, J., G. McClelland, M. McKee, W.D. Schulze, and E. Norden (1998), "Payoff dominance vs. cognitive transparency in decision making," *Economic Inquiry*, **36**(2), 272–85.

Jakobsson, K.M. and A.K. Dragun (2001), "The worth of a possum: Valuing species with the contingent valuation method," *Environmental and Resource Economics*, **19**(3), 211–27.

Jakus, J.P. (1992), "Valuing the private and public dimensions of a mixed good: An application to pest control," PhD thesis, North Carolina State University.

Jin, J., A. Indab, O. Nabangchang, T.D. Thuy, D. Harder, and R.F. Subade (2010), "Valuing marine turtle conservation: A cross-country study in Asian cities," *Ecological Economics*, **69**(10), 2020–26.

Johnston, R.J. and J.M. Duke (2009), "Informing preservation of multifunctional agriculture when primary research is unavailable: An application of meta-analysis," *American Journal of Agricultural Economics*, **91**(5), 1353–9.

Johnston, R.J., E.Y. Besedin, R. Iovanna, C.J. Miller, R.F. Wardwell, and M.H. Ranson (2005), "Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: A meta-analysis," *Canadian Journal of Agricultural Economics*, **53**(2–3), 221–48.

Kahneman, D. (1986), "Comments by Professor Daniel Kahneman," in R. Cummings, D. Brookshire, and W. Schulze (eds), *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, Totowa, NJ: Rowman & Allanheld, pp. 185–94.

Kahneman, D. and J.L. Knetsch (1992), "Valuing public goods: The purchase of moral satisfaction," *Journal of Environmental Economics and Management*, **22**(1), 57–70.

Kahneman, K. and I. Ritov (1994), "Determinants of stated willingness to pay for public goods: A study in the headline method," *Journal of Risk and Uncertainty*, **9**(1), 5–38.

Kling, C.L., D.J. Phaneuf, and J. Zhao (2012), "From Exxon to BP: Has some number become better than no number?," *Journal of Economic Perspectives*, **26**(4), 3–26.

Kniivila, M., V. Ovaskainen, and O. Saastamoinen (2002), "Costs and benefits of forest conservation: Regional and local comparisons in Eastern Finland," *Journal of Forest Economics*, **8**(2),131–50.

Kragt, M.E. (2013), "The effects of changing cost vectors on choices and scale heterogeneity," *Environmental and Resource Economics*, **54**(2), 201–21.

Krieger, D.J. (1994), "The economic value of environmental risk information: Theory and application to the Michigan sport fisher," dissertation, Michigan State University, Ann Arbor, MI: University Microfilms International.

Kwak, S., S. Yoo, and S. Han (2003), "Estimating public's value for urban forest in Seoul Metropolitan Area of Korea: A contingent valuation study," *Urban Studies*, **40**(11), 2,207–21.

Lindhjem, H. (2007), "20 years of stated preference valuation of non-timber benefits

from Fennoscandian forests: A meta-analysis," *Journal of Forest Economics*, **12**(4), 251–77.
Longo, A., D. Hoyos, and A. Markandya (2012), "Willingness to pay for ancillary benefits of climate change mitigation," *Environmental Resource Economics*, **51**(1), 119–40.
Loomis, J. and E. Ekstrand (1997), "Economic benefits of critical habitat for the Mexican spotted owl: A scope test using a multiple-bounded contingent valuation survey," *Journal of Agricultural and Resource Economics*, **22**(2), 356–66.
Loomis, J. and D.S. White (1996), "Economic benefits of rare and endangered species: Summary and meta-analysis," *Ecological Economics*, **18**(3), 197–206.
Loomis, J., A. González-Cabán and R. Gregory (1994), "Do reminders of substitutes and budget constraints influence contingent valuation estimates?," *Land Economics*, **70**(4), 499–506.
Loomis, J., L.T. Hung, and A. González-Cabán (2009), "Willingness to pay function for two fuel treatments to reduce wildfire acreage burned: A scope test and comparison of white and Hispanic households," *Forest Policy and Economics*, **11**(3), 155–60.
Loomis, J., M. Lockwood, and T. DeLacy (1993), "Some empirical evidence on embedding effects in contingent valuation of forest protection," *Journal of Environmental Economics and Management*, **24**, 44–55.
Macmillan, D.C. and E.I. Duff (1998), "Estimating the non-market costs and benefits of native woodland restoration using the contingent valuation method," *Forestry*, **71**(3), 247–59.
Macmillan, D., N. Hanley, and S. Buckland (1996), "A contingent valuation study of uncertain environmental gains," *Scottish Journal of Political Economy*, **43**(5), 519–33.
Magnussen, K. (1992), "Valuation of reduced water pollution using the contingent valuation method – testing for amenity misspecification," in S. Navrud (ed.), *Pricing the European Environment*, Oxford: Oxford University Press.
Markowski, M.A., K.J. Boyle, R.C. Bishop, D.M. Larson, and R.W. Paterson (2001), 'A cautionary note on interpreting meta analyses', unpublished paper, Industrial Economics Inc.
McDaniels, T.L., R. Gregory, J. Arvai, and R. Chuenpagdee (2003), "Decision structuring to alleviate embedding in environmental valuation," *Ecological Economics*, **46**(1), 33–46.
McFadden, D. (1994), "Contingent valuation and social choice," *American Journal of Agricultural Economics*, **76**(4), 689–708.
McFadden, D. and G. Leonard (1993), "Issues in the contingent valuation of environmental goods: Methodologies for data collection and analysis," in J.A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*, Amsterdam: Elsevier, pp. 177–207.
Metcalfe, P.J. (2012), "Non-market valuation using stated preferences: Applications in the water sector," thesis submitted to the Department of Geography and Environment, the London School of Economics and Political Science.
Mill, G.A, T.M. van Rensburg, S. Hynes, and C. Dooley (2007), "Preferences for multiple use forest management in Ireland: Citizen and consumer perspectives," *Ecological Economics*, **60**(3), 642–53.
Mørkbak, M.R., T. Christensen, and D. Gyrd-Hansen (2010), "Choke price bias in choice experiments," *Environmental Resource Economics*, **45**(4), 537–51.

Mullarkey, D.J. (1997), "Contingent valuation of wetlands: Testing sensitivity to scope," dissertation submitted to the Graduate School of the University of Wisconsin-Madison.
Mullarkey, D.J. and R.C. Bishop (1999), "Sensitivity to scope: Evidence from a CVM study of wetlands," presented at the American Agricultural Economics Association Annual Meeting.
Navrud, S. (1989), "Estimating social benefits of environmental improvements from reduced acid rain deposition: A contingent valuation survey," in H. Folmer and E.C. van Ierland (eds), *Valuation Methods and Policy Making in Environmental Economics*, Amsterdam: Elsevier.
Nunes, P.A.L.D. and E. Schokkaert (2003), "Identifying the warm glow effect in contingent valuation," *Journal of Environmental Economics and Management*, **45**(2), 231–45.
Ojea, E. and M.L. Loureiro (2011), "Identifying the scope effect on a meta-analysis of biodiversity studies," *Resource and Energy Economics*, **33**(3), 706–24.
Ojea, E., P.A.L.D. Nunes, and M.L. Loureiro (2010), "Mapping biodiversity indicators and assessing biodiversity values in global forests," *Environmental and Resource Economics*, **47**(3), 329–47.
Pattison, J., P.C. Boxall, and W.L. Adamowicz (2011), "The economic benefit of wetland retention and restoration in Manitoba," *Canadian Journal of Agricultural Economics*, **59**(2), 223–44.
Petrolia, D.R. and T. Kim (2009), "What are barrier islands worth? Estimates of willingness to pay for restoration," *Marine Resource Economics*, **24**(2), 131–46.
Poe, G.L., K.J. Boyle, and J.C. Bergstrom (2001), "A preliminary meta analysis of contingent values for ground water quality revisited," in J.C. Bergstrom, K.J. Boyle, and G.L. Poe (eds), *The Economic Value of Water Quality*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.
Poe, G., K.L. Giraud, and J.B. Loomis (2005), "Computational methods for measuring the difference of empirical distributions," *American Journal of Agricultural Economics*, **87**(2), 353–65.
Pouta, E. (2003), "Attitude-behavior framework in contingent valuation of forest conservation," dissertation submitted to the University of Helsinki.
Pouta, E. (2005), "Sensitivity to scope of environmental regulation in contingent valuation of forest cutting practices in Finland," *Forest Policy and Economics*, **7**(4), 539–50.
Powe, N.A. and I.J. Bateman (2004), "Investigating insensitivity to scope: A split-sample test of perceived scheme realism," *Land Economic*, **80**(2), 258–71.
Prelec, D., J. Burrows, and P. Dixon (forthcoming), "Context sensitivity in stated preference experiments."
Rahmatian, M. (1986), "Extensions of the disaggregate bid experiment: Variations in framing," *Journal of Environmental Management*, **22**(3), 191–202.
Randall, A., J.P. Hoehn, and G.S. Tolley (1981), "The structure of contingent markets: Some results of a recent experiment," paper presented at the annual meeting of the American Economic Association, Washington, DC.
Rathnayake, R.M.S. (2016), "Willingness to pay for a novel visitor experience: Ecotourism planning at Kawdulla National Park in Sri Lanka," *Tourism Planning and Development Journal*, **13**(1), 37–51.
Ready, R.C., M.C. Berger, and G.C. Blomquist (1997), "Measuring amenity

benefits from farmland: Hedonic pricing vs. contingent valuation," *Growth and Change*, **28**(4), 438–58.
Ressurreição, A., J. Gibbons, T.P. Dentinho, M. Kaiser, R.S. Santos, and G. Edwards-Jones (2011), "Economic valuation of species loss in the open sea," *Ecological Economics*, **70**(4), 729–39.
Richardson, L. and J. Loomis (2009), "The total economic value of threatened, endangered and rare species: An updated meta-analysis," *Ecological Economics*, **68**(5), 1535–48.
Rollins, K. and A. Lyke (1998), "The case for diminishing marginal existence values," *Journal of Environmental Economics and Management*, **36**(3), 324–44.
Rosenberger, R.S. and J.B. Loomis (2000a), *Benefit Transfer of Outdoor Recreation Use Values: A Technical Document Supporting the Forest Service Strategic Plan (2000 Revision)*, final report for the USDA Forest Service.
Rosenberger, R.S. and J.B. Loomis (2000b), "Using meta-analysis for benefit transfer: In-sample convergent validity tests of an outdoor recreation database," *Water Resources Research*, **36**(4), 1097–107.
Rowe, R.D., W.D. Shaw, and W. Schulze (1992), "Nestucca oil spill," in K.M. Ward and J.W. Duffield (eds), *Natural Resource Damages: Law and Economics*, New York: John Wiley, pp. 527–54.
Ryan, M. and S. Wordsworth (2000), "Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments," *Scottish Journal of Political Economy*, **47**(5), 504–24.
Samples, K.C. and J.R. Hollyer (1990), "Contingent valuation of wildlife resources in the presence of substitutes and complements," in R.L. Johnson and G.V. Johnson (eds), *Economic Valuation of Natural Resources: Issues, Theory, and Applications*, Boulder, CO: Westview Press, pp. 177–92.
Schkade, D.A. and J.W. Payne (1993), "Where do the numbers come from? How people respond to contingent valuation questions," in J.A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*, Amsterdam: Elsevier, pp. 271–93.
Schkade, D.A. and J.W. Payne (1994), "How people respond to contingent valuation questions: A verbal protocol analysis of willingness to pay for an environmental regulation," *Journal of Environmental Economics and Management*, **26**(1), 88–109.
Schulze, W.D., R.D. Rowe, and G.H. McClelland (1995), *Contingent Valuation of Natural Resource Damages Due to Injuries to the Upper Clark Fork River Basin*, report prepared for the State of Montana Natural Resource Damage Litigation Program.
Shrestha, R.K. and J.B. Loomis (2003), "Meta-analytic benefit transfer of outdoor recreation economic values: Testing out of sample convergent validity," *Environmental and Resource Economics*, **25**(1), 79–100.
Smith, A.E., M.A. Kemp, T.H. Savage, and C.L. Taylor (2005), "Methods and results from a new survey of values for eastern regional haze improvements," *Journal of Air and Waste Management Association*, **55**(11), 1767–79.
Smith, K.V. and L.L. Osborne (1996), "Do contingent valuation estimates pass a 'scope' test? A meta-analysis," *Journal of Environmental Economics and Management*, **31**(3), 287–301.
Smith, V.K. and Y. Kaoru (1990), "Signals or noise? Explaining the variation in recreation benefit estimates," *American Journal of Agricultural Economics*, **72**(2), 419–33.
Smith, V.K., X. Zhang, and R.B. Palmquist (1997), "Marine debris, beach

quality, and non-market values," *Environmental and Resource Economics*, **10**(3), 223–47.
Stanley, D.L. (2005), "Local perception of public goods: Recent assessments of willingness-to-pay for endangered species," *Contemporary Economic Policy*, **23**(2), 165–79.
Stevens, T.H., S. Benin, and J.S. Larson (1995), "Public attitudes and values for wetland conservation in New England," *Wetlands*, **15**(3), 226–31.
Stevens, T.H., N.E. DeCoteau, and C.E. Willis (1997), "Sensitivity of contingent valuation to alternative payment schedules," *Land Economics*, **73**(1), 140–48.
Streever, W.J., M. Callaghan-Perry, A. Searles, T. Stevens, and P. Svoboda (1998), "Public attitudes and values for wetland conservation in New South Wales, Australia," *Journal of Environmental Management*, **54**(1), 1–14.
Sturtevant, L.A., F.R. Johnson, and W.H. Desvousges (1998), "A meta-analysis of recreational fishing," unpublished manuscript, Durham, NC: Triangle Economics Research.
Svedsäter, H. (2000), "Contingent valuation of global environmental resources: Test of perfect and regular embedding," *Journal of Economic Psychology*, **21**(6), 605–23.
Tanguay, M., W. Adamowicz, P. Boxall, W. Phillips, and W. White (1993), *A Socio-Economic Evaluation of Woodland Caribou in Northwestern Saskatchewan*, Project Report No. 93-04.
Total Value Team (2015a), "Draft Technical Memo TM-10: Econometric analysis of choice questions," in the US Department of the Interior's Deepwater Horizon Response & Restoration Administrative Record.
Total Value Team (2015b), "Draft Technical Memo TM-5: Development and testing of the survey questionnaire," in the US Department of the Interior's Deepwater Horizon Response & Restoration Administrative Record.
Van Houtven, G., J. Powers, and S.K. Pattanayak (2007), "Valuing water quality improvements in the United States using meta-analysis: Is the glass half-full or half-empty for national policy analysis?," *Resource and Energy Economics*, **29**(3), 206–14.
Veisten, K., H.F. Hoen, S. Navrud, and J. Strand (2004a), "Scope insensitivity in contingent valuation of complex environmental amenities," *Journal of Environmental Management*, **73**(4), 317–31.
Veisten, K., H.F. Hoen, S. Navrud, and J. Strand (2004b), "Sequencing and the adding up property in contingent valuation of endangered species: Are contingent non-use value economic values?," *Environmental and Resource Economics*, **29**, 419–33.
Vo, D.T. and K.V. Huynh (2014), "Estimating residents' willingness to pay for groundwater protection in the Vietnamese Mekong Delta," *Applied Water Science*, November, DOI: 10.1007/s13201-014-0257-8.
Walsh, R.G., D.M. Johnson, and J.R. McKean (1992), "Benefit transfer of outdoor recreation demand studies, 1968–1988," *Water Resources Research*, **28**(3), 707–13.
Walsh, R., J. Loomis, and R. Gillman (1984), "Valuing option, existence, and bequest demands for wilderness," *Land Economics*, **60**(1), 14–29.
Welsh, M.P., R.C. Bishop, R.M. Baumgartner, and M.L. Phillips (1994), *Pilot Test Non-use Value Study (draft final report), Glen Canyon Environmental Studies Technical Report*, Madison, WI: HBRS Inc.
Welsh, M.P., R.C. Bishop, M.L. Phillips, and R.M. Baumgartner (1995), *Glen*

Canyon Environmental Studies Non-Use Value Study, final report prepared for Glen Canyon Environmental Studies Non-Use Value Committee, Madison, WI: RCG/Hagler-Bailly Inc.
White, P.C.L., K.W. Gregory, P.J. Lindley, and G. Richards (1997), "Economic values of threatened mammals in Britain: A case study of the otter *Lutra lutra* and the water vole *Arvicola terrestris*," *Biological Conservation*, **82**(3), 345–54.
Whitehead, J.C. (1992), "Ex ante willingness to pay with supply and demand uncertainty: Implications for valuing a sea turtle protection programme," *Applied Economics*, **24**(9), 981–8.
Whitehead, J.C. (2016), "Plausible responsiveness to scope in contingent valuation," *Ecological Economics*, **128**, 17–22.
Whitehead, J.C. and T.L. Cherry (2007), "Willingness to pay for a green energy program: A comparison of ex-ante and ex-post hypothetical bias mitigation approaches," *Resource and Energy Economics*, **29**(4), 247–61.
Whitehead, J.C. and S. Finney (2003), "Willingness to pay for submerged maritime cultural resources," *Journal of Cultural Economics*, **27**(3), 231–40.
Whitehead, J.C., P.A. Groothius, and R. Southwick (2007), "Linking recreation demand and willingness to pay with the inclusive value: Valuation of Saginaw Bay coastal marsh," presented at the US Environmental Protection Agency Workshop (Valuation for Environmental Policy: Ecological Benefits Workshop).
Whitehead, J.C., T.C. Haab, and J.-C. Huang (1998), "Part-whole bias in contingent valuation: Will scope effects be detected with inexpensive survey methods?," *Southern Economic Journal*, **65**(1), 160–68.
Whitehead, J.C., P.A. Groothius, R. Southwick, and P. Foster-Turley (2009), "Measuring the economic benefits of Saginaw Bay coastal marsh with revealed and stated preference methods," *Journal of Great Lakes Research*, **35**(3), 430–37.
Wilson, M.A. (2000), 'Rethinking scope sensitivity and contingent valuation surveys: Strong environmental attitudes and contingent economic values', dissertation submitted to the Graduate School of the University of Wisconsin-Madison, UMI Number: 9981890.
Woodward, R.T. and Y. Wui (2001), "The economic value of wetland services: A meta-analysis," *Ecological Economics*, **37**(2), 257–70.
Wu, P. (1993), "Substitution and complementarity in commodity space: Benefit evaluation of multidimensional environmental policy," *Academia Economic Papers*, **21**(1), 151–82.
Zandersen, M. and R.S.J. Tol (2009), "A meta-analysis of forest recreation values in Europe," *Journal of Forest Economics*, **15**(1–2), 109–30.