Valuing the Environment Through Contingent Valuation
Author(s): W. Michael Hanemann
Source: *The Journal of Economic Perspectives*, Vol. 8, No. 4 (Autumn, 1994), pp. 19-43
Published by: American Economic Association
Stable URL: http://www.jstor.org/stable/2138337
Accessed: 17-08-2017 19:48 UTC

Journal of Economic Perspectives—Volume 8, Number 4—Fall 1994—Pages 19–43

Valuing the Environment Through Contingent Valuation

W. Michael Hanemann

The ability to place a monetary value on the consequences of pollution discharges is a cornerstone of the economic approach to the environment. If this cannot be done, it undercuts the use of economic principles, whether to determine the optimal level of pollution or to implement this via Pigouvian taxes or Coase-style liability rules. Sometimes, the valuation involves a straightforward application of methods for valuing market commodities, as when sparks from a passing train set fire to a wheat field. Often, however, the valuation is more difficult. Outcomes such as reducing the risk of human illness or death, maintaining populations of native fish in an estuary, or protecting visibility at national parks are not themselves goods that are bought and sold in a market. Yet, placing a monetary value on them can be essential for sound policy.

The lack of a market to generate prices for such outcomes is no accident. Markets are often missing in such cases because of the nonexcludable or nonrival nature of the damages: for those affected by it, pollution may be a public good (or bad). The public good nature of the damages from pollution has several consequences. It explains, for example, why the damages are sometimes large—only a few people may want to own a sea otter pelt, say, but many may want this animal protected in the wild. It also explains why market prices are inappropriate measures of value. In the presence of externalities, market transactions do not fully capture preferences. Collective choice is the more relevant paradigm.

This is precisely what Ciriacy-Wantrup (1947) had in mind when he first proposed the contingent valuation method. Individuals should be interviewed

■ *W. Michael Hanemann is Associate Professor of Agricultural and Resource Economics, University of California, Berkeley, California.*

and "asked how much money they are willing to pay for successive additional quantities of a collective extra-market good." If the individual values are aggregated, "the result corresponds to a market-demand schedule" (p. 1189). Thus, surveys offered a way to trace the demand curve for a public good that could not otherwise be gleaned from market data. Schelling (1968) made a similar point in his paper on valuing health. While the price system is one way to find out what things are worth to people, he wrote, another way is to ask people, whether through surveys or votes. Answering surveys may be hypothetical, but no more than buying unfamiliar or infrequent commodities. "In any case, relying exclusively on market valuations and denying the value of direct enquiry in the determination of government programs would depend on there being for every potential government service, a close substitute available in the market at a comparable price. It would be hard to deduce from first principles that this is bound to be the case" (pp. 143–4).

Schelling's point was not that indirect methods using market transactions have no role, but rather that they cannot always be counted on to provide a complete measure of value. Analysts can often capture some effects of a change in air quality or a change in risk to human health through a hedonic analysis that looks for evidence to property values or wage rates (Rosen, 1974). But people may also value those items in ways not reflected in wages or property values. Similarly with averting expenditures and household production models (Freeman, 1993), which rely on the demand for market commodities that are complements to, or surrogates for, the nonmarket good. If people value that good at least partly for reasons unrelated to their consumption of the complementary private goods, those methods capture just part of people's value—what is called the "use value" component, following Krutilla (1967).[1] They fail to measure the "non-use value" or "existence value" value component, which contingent valuation can capture.

An alternative is to turn to the political system, for example using collective choice models to estimate demands for local public goods (Oates, 1994). However, Cropper (1994) suggests this is unlikely to be useful for the environment because, in the United States, there are few cases where local governments actually set environmental quality. Moreover, as Chase (1968) noted, the method contains an element of circularity: a major reason for the spread of benefit-cost analysis is legislators' desire to obtain information on the public's value for government programs. While it may sometimes be desirable to leave the assessment of value to the legislative process, it is not obvious that this is always so. Measuring liability for damages from pollution is an example. In some cases one wants to ascertain how the public values something, and contingent valuation may be the only way to measure this short of a plebiscite.

Ciriacy-Wantrup (1947) recognized that surveys are not foolproof. The degree of success depends on the skill with which the survey is designed and

[1]For a formal definition, see Hanemann (1994a).

implemented. But it was time, he felt, that economics took advantage of developments in social psychology and the newly emerging academic field of survey research: "Welfare economics could be put on a more realistic foundation if a closer cooperation between economics and certain young branches of applied psychology could be established" (p. 1190). This finally occurred in the 1980s, and contingent valuation came of age. Two landmarks were an EPA conference in 1984 that brought together leading practitioners, other economists, and psychologists to assess the state-of-the-art (Cummings et al., 1986), and the publication of what has become the standard reference on contingent valuation, Mitchell and Carson (1989), which puts it in a broader context involving elements from economics, psychology, sociology, political science, and market research.

Contingent valuation is now used around the world (Navrud, 1992; Bateman and Willis, forthcoming), both by governments agencies and the World Bank for assessing a variety of investments. A recent bibliography lists 1600 studies and papers from over 40 countries on many topics, including transportation, sanitation, health, the arts and education, as well as the environment (Carson et al., 1994c). Some notable examples are Randall, Ives and Eastman (1974) on air quality in the Four Corners area, the first major non-use value study; Brookshire et al. (1982) on air pollution in Southern California; Carson and Mitchell (1993) on national water quality benefits from the Clean Water Act; Smith and Desvousges (1986) on cleaning up the Monongahela River, Jones-Lee, Hammerton and Phillips (1985) on highway safety; Boyle, Welsh and Bishop (1993) on rafting in the Grand Canyon; Briscoe et al. (1990) on drinking water supply in Brazil; and the study on the *Exxon Valdez* oil spill I helped conduct for the State of Alaska (Carson et al., 1992).

This paper focuses generally on the use of contingent valuation to measure people's values for environmental resources, rather than specifically on natural resource damages. It will describe how researchers go about conducting reliable surveys. It then addresses some common objections to surveys and, lastly, considers the compatibility between contingent valuation and economic theory.

Conducting Reliable Surveys

In all research, details matter. How a contingent valuation survey is conducted is crucial. While there is no panacea, various procedures have been developed in recent years that enhance the credibility of a survey and make it more likely to produce reliable results. These touch all aspects, including sampling, instrument development, formulation of the valuation scenario, questionnaire structure, and data analysis. The main ways of assuring reliability are summarized here.

Suppose one approached people in a shopping mall, made them put their bags down for a moment, and asked them what was the most they would be

willing to pay for a sea otter in Alaska or an expanse of wilderness in Montana. This is how the President of American Petroleum Institute and other critics have characterized contingent valuation (DiBona, 1992). The essence of their argument is summarized in titles such as "Ask a Silly Question" and "Pick a Number" (Anon., 1991; Bate, 1994). It does not require any unusual perspicacity to see that this approach is unlikely to produce reliable results. For precisely this reason, it is *not* what good contingent valuation researchers do, and it is *not* what was recommended by the NOAA Panel on Contingent Valuation (Arrow et al., 1993) described in Portney's paper in this issue.

Serious surveys of the general public avoid convenience sampling, such as stopping people in the street; they employ statistically based probability sampling.[2] They also avoid self-administered surveys, such as mail surveys or questionnaires handed out in a mall, because of the lack of control over the interview process. For a major study, the NOAA Panel recommended in-person interviews for their superior reliability. Furthermore, interviews should occur in a setting that permits respondents to reflect and give a considered opinion, such as their home. Unless the study deals with consumer products, shopping malls are a poor choice. Indeed, the only contingent valuation study where people were stopped for a few minutes in a mall was one performed for Exxon (Desvousges et al., 1992).

The crux is how one elicits value. The two key developments have been to confront subjects with a specific and realistic situation rather than an abstraction, and to use a closed-ended question which frames the valuation as voting in a referendum.

A common temptation is to characterize the object of valuation in rather general terms: "What would you pay for environmental safety?" "What would you pay for wilderness?" The problem is that these are abstractions. People's preferences are not measured in the abstract but in terms of specific items. "Paying for wilderness" is meaningless; what is meaningful is paying higher taxes or prices to finance particular actions by somebody to protect a particular wilderness in some particular manner. Therefore, one wants to confront respondents with something concrete. Moreover, one should try to avoid using counterfactuals. "What would you pay not to have had the *Exxon Valdez* oil spill?" is utterly hypothetical because one cannot undo the past. By contrast, "What would you pay for this new program that will limit damage from any future oil spills in Prince William Sound?" offers something that is tangible.

The goal in designing a contingent valuation survey is to formulate it around a specific commodity that captures what one seeks to value, yet is plausible and meaningful. The scenario for providing the commodity may be real; if not, the key is to make it seem real to respondents. They are not actually

[2]DiBona's scenario actually was the practice in the 1930s when most surveys were "brief encounters" on the street or in stores (Smith, 1987). The 1940s saw the adoption of probability sampling, standardized survey techniques, longer and more complex survey instruments, and in-depth focused interviews (Merton and Kendall, 1946).

making a payment during the interview, but they are expressing their intention to pay. The vaguer and less specific the commodity and payment mechanism, the more likely respondents are to treat the valuation as symbolic. To make the payment plausible, one needs to specify the details and tie them to provision of the commodity so this cannot occur without payment. There should be a clear sense of commitment; for example, if the program is approved, firms will raise prices, or the government taxes, so there is no avoiding payment once a decision is made.[3]

Until the mid-1980s, most contingent valuation surveys used some version of an open-ended question, like "What is the most you would be willing to pay for . . . ?" Since then, most major contingent valuation studies have used closed-ended questions like "If it cost $\$x$, would you be willing to pay this amount?" or "If it cost $\$x$, would you vote for this?" Different people are confronted with different dollar amounts. Plotting the proportion of "yes" responses against the dollar amount traces out the cumulative distribution function of willingness-to-pay.[4]

Of course, if people carried utility functions engraved in their brains, the question format would not matter. But they don't, and it does matter. In this country, posted prices are the norm rather than bargaining. In market transactions people usually face discrete choices: here is an item, it costs $\$x$, will you take it? Similarly in voting. Moreover, there is abundant evidence that respondents find the open-ended willingness-to-pay question much more difficult to answer than the closed-ended one; for market and nonmarket goods alike, people can generally tell you whether they would pay some particular amount, but they find it much harder to know what is the *most* that they would possibly pay. Indeed, the experience with open-ended willingness-to-pay questions for market goods is that people are more likely to tell you what the good costs than what it is worth to them. In addition to being less realistic and harder to answer, the open-ended format creates incentives which are different from those in the closed-ended format. With the open-ended format, as with an oral auction, there are strategic reasons for stating less than one's full value—a theoretical result strongly supported by experimental evidence. This is not so with a closed-ended format; there, the NOAA Panel held, there is no strategic reason for the respondent to do other than answer truthfully.[5]

For these reasons, the NOAA Panel considered the closed-ended format combined, where possible, with a voting context the most desirable for

[3]To underscore this, the interviewer may tell respondents that the government uses surveys like this to find out whether taxpayers are willing to pay for new programs it is considering.

[4]The methodology here is to assume a random utility model for individual preferences. This can be estimated using standard techniques for binary choices. Bishop and Heberlein (1979) were the first to use this format; the link with utility theory was developed in Hanemann (1984).

[5]With auctions, it is well documented that formal matters and that oral auctions generate lower prices than posted-price auctions. Why the surprise when the same holds true for open- versus closed-ended payment questions?

contingent valuation: "The simplest way to approach the valuation problem," it held, "is to consider a contingent valuation survey as essentially a self-contained referendum in which respondents vote to tax themselves for a particular purpose" (p. 20). This is a rather different conception of contingent valuation from asking silly questions of passers-by.

In his introduction to this symposium, Portney describes other ways to make a contingent valuation questionnaire more reliable: providing adequate and accurate information; making the survey balanced and impartial; insulating it from any general dislike of big business; reminding respondents of the availability of substitutes, and of their budget constraint; facilitating "don't know" responses; allowing respondents to reconsider at the end of the interview. Several steps can be taken to eliminate any perception of interviewer pressure. At the outset, the interviewer can assure respondents that there are no "right" answers. Before asking the voting question, to legitimate a negative response, the interviewer could say something like: "We have found that some people vote for the program and others vote against. Both have good reasons for voting that way," and then list some reasons for saying "no."[6] Another possibility is if the interviewer does not actually see the respondents' votes, for example by having them write on a ballot placed in a sealed box.

A recent innovation, considered essential by the NOAA Panel, is a "debriefing" section at the end of the survey. This checks respondents' understanding and acceptance of key parts of the contingent valuation scenario. For example, was the damage as bad as described? Did you think the program would work? Did you think you really would have to pay higher taxes if the program went through? This also probes the motives for their answer to the willingness-to-pay question. What was it about the program that made you decide to vote for it? Why did you vote no? Moreover, throughout the survey, all spontaneous remarks by the respondent are recorded verbatim as they occur. After the survey, the interviewer is debriefed and asked about the circumstances of the interview, how attentive the respondent was, whether the respondent seemed to understand the questions and appeared confident in his responses. In this way, one creates a rich portrait of the interview. This information can be exploited in the data analysis. One can monitor for the misunderstandings, measure statistically how they affected respondents' willingness-to-pay, and adjust accordingly. For example, if a subject who voted "yes" appeared to be valuing something different than the survey intended, this case can be dropped or the "yes" converted to a "no."

With any data, different statistical procedures can produce different results. The closed-ended format raises several statistical issue, for example, one might summarize the willingness-to-pay distribution by using its mean, or its

[6]For example, the interviewer might note that some people prefer to spend the money on other social or environmental problems instead, or they find the cost is more than they can afford or than the program is worth, or they cannot support the program because it would benefit only one area (Carson et al., 1992).

median, or another quantile. The mean is extremely sensitive to the right tail of the distribution; that is, to the responses of the higher bidders. For this reason, if the mean is to be used, a nonparametric or bounded influence approach is highly recommended for fitting the willingness-to-pay distribution. The median, by contrast, is usually very robust (Hanemann, 1984). Another issue is that the choice of dollar bids affects the precision with which the parameters of the willingness-to-pay distribution are estimated; significant improvements can be achieved by using optimal experimental designs (Kanninen, 1993). Statistical techniques can also be used to probe for yea-saying or other response effects, and correct for them if they are present (Hanemann and Kanninen, forthcoming).

While none of these alone is decisive, taken together they are likely to produce a reliable measure of value. Apart from the expense of in-person interviews, they are all eminently feasible.[7] Other essential ingredients are relentless attention to detail and rigorous testing of the instrument, usually in collaboration with survey experts, so that the researcher understands exactly how it works in the field and is sure it communicates what was intended.

It is no coincidence that the handful of studies that Diamond and Hausman select from the contingent valuation literature in their companion paper in this issue violate most of these precepts, as do the Exxon surveys reported in Hausman (1993). None uses in-person interviews. Many are self-administered. Most use open-ended questions. None is cast as voting.[8] Many ask questions with a remarkable lack of detail.[9] Several seem designed to highlight the symbolic aspects of valuation at the expense of substance.[10] The Exxon surveys were designed and fielded in great haste, with little pretesting, just at a time when federal agencies were gearing up for natural resource damage regulations.[11] The only way to justify this is to make the tacit assumption that, if

[7]Is there an acceptable alternative to in-person surveys? The NOAA Panel felt mail surveys have significant problems rendering them unsuitable. Telephone surveys avoid these problems, but preclude the use of visual aids and need to be short. The most promising alternative is a mail/telephone combination in which an information package is mailed to respondents who are then interviewed by phone (Hanemann, Loomis and Kanninen, 1991). This permits an extensive phone interview which seems to provide many of the benefits of an in-person survey at much lower cost.

[8]Two studies Diamond and Hausman cite as showing a lack of commitment in contingent valuation, Seip and Strand (1992) and Duffield and Patterson (1991), used open-ended questions about payment to an environmental charity. Most of Seip and Strand's subjects who were followed up afterwards said that they had been expressing their willingness-to-pay for environmental problems generally, rather than the particular environmental group. Careful pretesting would have discovered this beforehand.

[9]This is notably a problem in Diamond et al. (1993).

[10]Including Kahneman and Ritov (1993), Kahneman and Knetsch (1992), and Kemp and Maxwell (1993). The last two employ a "top-down" procedure in which respondents are given details of the item only *after* they value it. They are first confronted with something broad, like "preparedness for disasters." After stating their willingness-to-pay for the broad category, they are told what it comprises and asked their willingness-to-pay for *one* of those components. Then, they are told what *this* comprises, and so on. The *change* in the *quantity* of any item is never specified.

[11]Hanemann (1994a, b) critiques these studies.

contingent valuation is valid, details of its implementation should not matter. This is fundamentally wrong: measurement results are not invariant with respect to measurement practice in *any* science.

Objections to Surveys

McCloskey (1985, p. 181) observes that economists generally dislike surveys: "Economists are so impressed by the confusions that might possibly arise from questionnaires that they have turned away from them entirely, and prefer the confusions resulting from external observation." In this section, I discuss four common objections to surveys.

Surveys are Vulnerable to Response Effects

Small changes in question wording or order sometimes cause significant changes in survey responses (Schuman and Presser, 1981). Since virtually all data used in economics come from surveys (including experiments, which are a form of survey), and all surveys are vulnerable to response effects, it is important to understand why these arise and how they can be controlled. A consensus is beginning to emerge based on insights from psychology and linguistics. Answering survey questions requires some effort, usually for no apparent reward. Respondents must interpret the meaning of the question, search their memory for pertinent information, integrate this into a judgment, and communicate the judgment to the interviewer. Although some are motivated to make the effort, others may become impatient, disinterested, or tired. Instead of searching for an accurate and comprehensive answer, they satisfice, just aiming for some response that will be accepted. Furthermore, interviews are interactions governed by social and linguistic norms that shape assumptions and expectations. Viewing respondents as satisficing agents following norms of conversation has proved helpful in interpreting survey data, explaining response effects, and designing more effective surveys (Groves, 1989; Krosnick, 1991).

Not all response phenomena are equally intractable. Some, such as order effects (for example, bias towards the first item in a list), can be detected and controlled, either by choosing the sequence that produces a conservative result or by randomizing the order of items across interviews.

A second type of effect is where there is a shift in meaning. This is substance, not noise. For example, similar words turn out to mean different things: "allow" is not the same as "not forbid," nor "higher prices" the same as "higher taxes."[12] Or there are framing effects, where subjects respond differ-

[12]And different words can mean the same thing, as in the movie *Annie Hall* where Woody Allen and Diane Keaton are asked by their psychiatrists how often they have sex. He says: "Hardly ever, maybe three times a week." She says: "Constantly, I'd say three times a week." With consumer expenditure surveys, Miller and Guin (1990) attest that life imitates art.

ently to situations the researcher saw as equivalent. It has been shown through debriefings that the subjects perceived the situations as substantively different, because either the researcher induced an unintended change in meaning or context, or the subjects made inferences that went beyond the information given (Frisch, 1993).[13] In each case, the shift in meaning is a source of error only if the researcher is unaware of it. Through rigorous testing with cognitive techniques, the researcher can come to understand exactly what the instrument means to people, and what they mean in response.[14]

A third phenomenon arises from the inherent difficulty of the task assigned the respondent. In recalling past events or behavior, for example, respondents resort to rounding, telescoping (time compression) and other inferential strategies that yield inaccurate reports of magnitudes and frequencies.[15] Bradburn et al. (1987) emphasize that factual and attitudinal surveys share many similar cognitive processes and errors. There is no easy solution for recall errors. This continues to be a problem for many data used by economists,[16] though not for contingent valuation data since there is no recall.

One cannot avoid the fact that surveys, like all communication, are sensitive to nuance and context and are bound by constraints of human cognition. One tries to detect discrepancies and repair them, but they cannot be entirely ruled out. It is important to keep a sense of proportion. As far as I know, nobody has stopped using data from the Current Population Survey, Consumer Expenditure Survey, Monthly Labor Survey, or Panel Study on Income Dynamics because there are response effects in such surveys. The same should apply to contingent valuation surveys.

The Survey Process Creates the Values

It has been asserted that contingent valuation respondents have no real value for the item, but just make one up during the course of the interview: the process creates the values that it seeks to measure. Debriefings can identify

[13]When there is incomplete information in a survey, respondents may go ahead and make their own assumptions. Consequently, the researcher loses control over his instrument. Diamond et al. (1993) is a contingent valuation example.

[14]On testing by federal survey agencies, see Tanur (1992). Lack of adequate testing can explain some notable violations of procedural invariance—respondents saw cues or meaning which the researcher didn't intend and failed to detect. An example is the base rate fallacy where "when no specific information was given, prior probabilities are properly utilized; when worthless evidence is given prior probabilities are ignored" (Tversky and Kahneman, 1974). A norm of conversation is to present information one believes relevant. That this was the expectation of subjects could have been detected through debriefings. On violations of conversational norms in base-rate experiments, see Krosnick, Li and Lehman (1990).

[15]Some pronounced telescoping errors are to be found in the Alaska recreation survey conducted by Hausman, Leonard and McFadden (1993).

[16]Juster and Stafford (1991) and Mathiowetz and Duncan (1988) discuss biases in labor supply estimates due to problems with bunching and misreporting in Current Population Survey data. Atkinson and Micklewright (1983) discuss errors in Family Expenditure Survey reports of income and its components. Other inconsistencies between micro- and macro-data sets for the household sector are discussed in Maki and Nishiyama (1993).

whether subjects were inattentive or unfocused and offered hasty or ill-considered responses, and these can be discarded if desired. But, the issue raised here is more fundamental. Diamond and Hausman feel they know real preferences when they see them, and they do not see them in contingent valuation. Based on the debriefing statements in Schkade and Payne (1993) that show most subjects, faced with an open-ended willingness-to-pay question, think about either what the item could cost or what they have spent on something remotely similar, Diamond and Hausman conclude that these people are just making up their answer rather than evincing "true economic preferences." But, what are "true economic preferences"? If a subject responds thoughtfully to a question about voting to raise taxes for a public good, by what criterion is that not a valid preference?

It is true that economists often assume consumer choice reflects an individual's global evaluation of alternatives, a "top-down" or "stored-rule" decision process. The stored-rule notion traces back to Hobbes and the English empiricists who conceived of cognition in terms of storing and retrieving "slightly faded copies of sensory experiences" (Neisser, 1967). Wilson and Hodges (1992) call this the "filing cabinet" concept of the mind. It long dominated not only economics but also psychology. But it is now being abandoned in the face of accumulating evidence from the neurosciences (Rose, 1992) and elsewhere that all cognition is a constructive process—people construct their memories, their attitudes, and their judgments. The manner of construction varies with the person, the item, and the context. A general principle is that people are cognitive misers: they tend to resolve problems of reasoning and choice in the simplest way possible. This is the emerging consensus not only in survey research, but also in social psychology, political psychology, and market research (Martin and Tesser, 1992; Sniderman, Brody and Tetlock, 1991; Payne, Bettman and Johnson, 1988).

For non-habituated and complex consumer choices, people often make "bottom-up" decisions; that is, they make up a decision rule at the moment they need to use it (Bettman, 1988). Olshavsky and Granbois (1979, p. 98) found that "for many purchases a decision process never exists, not even on first purchase." Bettman and Zins (1977) found that grocery shoppers construct a choice heuristic "on the spot" about 25 percent of the time; bottom-up construction of preferences occurred especially for meat and produce "as might be expected, since consumers cannot really rely on brand name for most choices of this type," less often for beverages and dairy products "where either strong taste preferences may exist or only a limited number of brands are available" (p. 81). This calls to mind a remark by Robert Solow that the debriefings in Schkade and Payne "sound an awful lot like Bob Solow in the grocery store." I suppose critics of contingent valuation would consider that Solow does not have true economic preferences, or that he has true economic preferences when buying milk but not meat.

The real issue is not whether preferences are a construct but whether they are a *stable* construct. While this surely varies with circumstances, the evidence

for contingent valuation is quite strong. There is now a number of test-retest studies in the contingent valuation literature, and these show both consistency in value over time and a high correlation at the individual level (Carson et al., 1994b). These levels of consistency are comparable to the most stable social attitudes such as political party identification.

Ordinary People are Ill-Trained For Valuing the Environment

If, as the NOAA Panel suggests, the goal of a contingent valuation survey is to elicit people's preferences as if they were voting in a referendum, then prior experience or training are irrelevant. These are not a criterion for voting.[17] Nor is their absence an argument against contingent valuation per se. Through direct questioning, one can readily identify which respondents knew of the issue before the interview, or before the oil spill, and determine whether they hold different values from those who did not. How one proceeds in calculating aggregate willingness-to-pay is something that can be decided separately from the survey. Who has standing, and whose values should count, are questions that we as economists have no special competence to judge.

Survey Responses Can't Be Verified

There are three ways to validate contingent valuation results: replication, comparison with estimates from other sources, and comparison with actual behavior where this is possible. Replication is useful even on a small scale both to see if results hold up and to check whether the instrument is communicating as intended. This is the single best way for a researcher to determine whether somebody's survey instrument works as claimed.

When contingent valuation measures direct use values, it may be possible to make a comparison with estimates obtained through indirect methods. Knetsch and Davis (1966) conducted the first test, comparing contingent valuation and travel demand estimates (a method described in Portney's paper) of willingness-to-pay for recreation in the Maine woods. The difference was less than 3 percent. There are now over 80 studies, offering several hundred comparisons between contingent valuation and indirect methods. The results are often fairly close; overall, the contingent valuation estimates are slightly

[17]Voter ignorance is a constant refrain for Diamond and Hausman. They use it to form a syllogism: voters are ill-informed, contingent valuation is like a referendum, therefore contingent valuation respondents are ill-informed. Both parts are false. Contingent valuation researchers take pains to ensure their samples are representative and their questionnaires intelligible, informative, and impartial, thus avoiding the vagaries of turnout and biased advertising in election campaigns. This is why political scientists are becoming interested in "deliberative polling"—in effect, extended contingent valuation surveys (Fishkin, 1991). Many analysts see a substantial core of rationality in voter behavior. Cronin (1989) finds Magleby's (1984) assessment of voter ignorance in referenda overblown. Fiorina (1981) and McKelvey and Ordeshook (1986) emphasize how campaign protagonists use signals to inform voters. Lupia (1993) analyzes the insurance reform battle in the 1988 California ballot and finds that informational "short cuts" enabled poorly informed voters to act as though they were well informed. What Sniderman (1993) calls "the new look in public opinion research" stresses how ordinary citizens use the information at hand to make sense of politics.

lower than the revealed preference estimates and highly correlated with them (Carson et al., 1994a).

The ideal is direct testing of contingent valuation predictions against actual behavior. There are about ten such tests in the literature. Diamond and Hausman mention only five of these. The ones not mentioned yield results quite favorable to contingent valuation.

Bohm (1972) conducted the first test, where subjects in Stockholm were asked their willingness-to-pay to see a new TV program. In five treatments, the program was shown if the group raised 500 Kr, with actual payment based in various ways on stated willingness-to-pay. A sixth treatment asked subjects what was the highest amount they would have given *if* they had been asked to pay an individual admission fee. The mean response was 10.2 Kr (about $2) when the group was asked a hypothetical question, versus an overall average of 8.1 Kr when the group actually paid. The difference between contingent valuation and non-contingent valuation means was not statistically significant in four of the five cases.

Bishop and Heberlein (1990) conducted a series of experiments with hunters who had applied for a deer-hunting permit in a favored game preserve run by the state of Wisconsin. The most relevant for current practice is an experiment in which they wrote to two groups of hunters offering to sell them a permit at a specified price. In one case, this was a real offer; in the other, it was asked as a hypothetical question. Estimated willingness-to-pay was $31 in the real sale versus $35 in the hypothetical sale, a statistically insignificant difference.

Dickie, Fisher and Gerking (1987) offered boxes of strawberries door-to-door at different prices. One treatment was a real offer—the household could buy any number of boxes at this price. The other asked how many boxes they *would* buy if these were offered at the given price. The resulting two demand curves were not significantly different. The parameter estimates were actually more robust over alternative model specifications for the hypothetical than the actual data (Smith, 1994).

Carson, Hanemann and Mitchell (1986) tested the accuracy of voting intentions in a water quality bond election in California in 1985. Closed-ended contingent valuation questions were placed on the Field California Poll a month before the vote, using different figures for the household cost. Adjusted for "don't know" responses, the predicted proportion of yes votes at the actual cost was 70–75 percent. The ballot vote in favor was 73 percent.

Cummings, Harrison and Rutstrom (1993) offered subjects small commodities at various prices. For one group, it was a real sale. A second group was first asked a hypothetical contingent valuation question—this item is not actually for sale but, if it were, would you buy it now? The experimenter then announced that, after all, she *would* sell the item, but they should feel free to revise their answer. When juicers were the item, 11 percent actually bought them in the real sale; with the second treatment, 41 percent said they would buy it if it were on sale, but then only 16 percent did. The 41 percent and 11

percent are significantly different. With calculators, 21 percent would buy in the hypothetical sale, versus 8 percent in the real sale. One wonders whether some respondents interpreted the question as "*if you needed a juicer*, would you buy this one?" Smith (1994) shows that the calculator responses do not generate a downward sloping demand curve for either the actual or hypothetical data. The experimental procedure contained nothing to emphasize commitment or counteract yea-saying in the hypothetical treatment. Cummings and his colleagues have recently added wording like the "reasons to say no" mentioned earlier. In one case, this reduced the hypothetical yes for calculators from 21 percent to 10 percent, not significantly different from the real 8 percent; in another there was no effect (Cummings, 1994).

Other contingent valuation tests have used open-ended payment questions, with predictable difficulties. Boyce et al. (1989) measured willingness-to-pay and willingness-to-accept for a house plant, with mixed results; Neill et al. (1994) measured willingness-to-pay for a map and a picture, with negative results. Both confound the issue by comparing contingent valuation responses to an experimental auction, begging the question of whether auction behavior understates willingness-to-pay. Duffield and Patterson (1991) and Seip and Strand (1992) compare actual and hypothetical contributions to an environmental cause. Diamond and Hausman focus on these studies because they showed a significant difference. But, soliciting an intention to make a charitable donation is a poor test of contingent valuation, because it invites less commitment than soliciting an intention to vote for higher taxes. To make things worse, Seip and Strand used members of the environmental group as the interviewers in their hypothetical treatment, thus increasing pressures for compliance. They compared hypothetical phone responses with responses to an actual mail solicitation. Duffield and Patterson compared hypothetical mail solicitations from the University of Montana with actual mail solicitations from the Nature Conservancy. In both studies, the difference in survey administration introduces a confounding factor which undermines the comparison.[18]

A cleaner test is provided by Sinden (1988) who conducted a series of 17 parallel experiments soliciting actual and hypothetical monetary donations to a fund for assisting soil conservation or controlling eucalypt dieback. In all 17 cases, there was no statistical difference between actual and hypothetical willingness-to-pay.

Thus, there is some substantial evidence for the validity of contingent valuation survey responses, although more studies are certainly needed. Many existing studies do not incorporate the refinements in contingent valuation

[18]The problem with mail surveys is that people may think the survey is junk mail and throw it out unopened. Duffield and Patterson made no allowance for the difference in sponsor identity on the envelope, which could explain the difference in response rates (Schuman, 1992). Response rates apart, the pattern of contributions was similar in the two treatments. Seip and Strand made no allowance for the fact that phone and mail solicitations generally have different response rates. Infosino (1986) found a sales rate three times higher with telephone than mail in an AT & T marketing effort.

method, described earlier, that emphasize realism and commitment. In this respect, the test by Carson, Hanemann and Mitchell (1986) points in the right direction because it deals directly with expression of voting intentions. The positive results in that study are consistent with other evidence showing that polls in this country reliably indicate public sentiment at the time they are taken, and polls close to an election are generally accurate predictors of the outcome.[19] Kelley and Mirer (1974) found voting intentions correctly predicted the actual vote in four presidential elections for 83 percent of those respondents who voted.[20] Surveys of purchase intentions in market research may not be accurate predictors of subsequent purchase behavior, but surveys of voting intentions are.[21]

Contingent Valuation and Economic Theory

Critics of contingent valuation like Diamond and Hausman, and their coauthors in Hausman (1993), reject contingent valuation as a method of economic valuation because the results of contingent valuation studies are inconsistent with economic theory as they see it. These assertions have become quite widely known. However, careful examination shows that in some cases the claims are not supported by the findings in the contingent valuation literature, and in others they rest on unusual notions about what economic theory does or does not prescribe. I briefly review these issues here, leaving a more detailed treatment to Hanemann (1994a).

Diamond and Hausman, and Milgrom (1993), make a number of statements about what is a permissible argument in a utility function. They argue that people should care about outcomes, not about the process whereby these are generated. People should not care whether animals are killed by man or die naturally. They should not care about details of provision or payment for a

[19]Diamond and Hausman seem troubled that voters change their minds during the course of an election campaign. They cite a 1976 electricity rate proposition in Massachusetts where support went from 71 percent in February to 25 percent in the November ballot. They fail to mention the reasons. Magleby (1984, p. 147) identifies opposition spending as the chief cause of such opinion reversals, and that certainly occurred in 1976—opponents outspent supporters more than threefold. In May, the Dukakis administration came out against it, as eventually did businesses, the unions, hospitals, colleges, and major newspapers.

[20]Ajzen and Fishbein (1980) offer some reasons to expect a high level of attitude-behavior correspondence for voting in terms of their theory of reasoned action.

[21]One reason for the difference is timing: unlike elections, people generally control the timing of their market purchases. The result is they may end up buying the commodity, but later than they said (Juster, 1964). This is especially likely for durables, the focus of much literature, since their durability permits delay in replacement. This is consistent with findings that purchase intentions are significantly more accurate for nondurables than durables (Ferber and Piskie, 1965); intentions *not* to purchase durables are highly accurate (Theil and Kosobud, 1968); and predictions of the brand selected when the purchase *does* occur tend to be highly accurate (Ajzen and Fishbein, 1980; Warshaw, 1980).

commodity, only price. Above all, they should value things for purely selfish motives. In their accompanying piece, Diamond and Hausman phrase this argument by saying that respondents should not contemplate "what they think is good for the country," because that reflects "warm glow" rather than "true economic preferences."[22] From this perspective, contingent valuation is unacceptable because it picks up existence values; for those to be allowed in a benefit-cost analysis, Milgrom (1993, p. 431) argues, "it would be necessary for people's individual existence values to reflect only their own personal economic motives and not altruistic motives, or sense of duty, or moral obligation."[23]

This criticism hardly comports with the standard view in economics that decisions about what people value should be left up to them. For example, Kenneth Arrow (1963, p. 17) wrote: "It need not be assumed here that an individual's attitude toward different social states is determined exclusively by the commodity bundles which accrue to his lot under each. The individual may order all social states by whatever standards he deems relevant." Or as Gary Becker (1993, p. 386) writes: "[I]ndividuals maximize welfare *as they conceive it*, whether they be selfish, altruistic, loyal, spiteful, or masochistic." When estimating demand functions for fish prior to Vatican II, no economist ever proposed removing Catholics because they were eating fish out of a sense of duty. Nor, when estimating collective choice models, do we exclude childless couples who vote for school bonds because they lack a personal economic motive.

A more substantive matter is how willingness-to-pay varies with factors that could reasonably be expected to influence it. This has been raised in connection with the embedding effect and the income elasticity of willingness-to-pay. Regarding the latter, Diamond and Hausman assert in this issue that the income effects measured in typical contingent valuation surveys are lower than would be expected if true preferences are measured. McFadden and Leonard (1993, p. 185) make the more specific claim that an income elasticity of willingness-to-pay less than unity constitutes grounds for doubting the validity of the contingent valuation method. There is no basis for either assertion. In the literature on the demand for state and local government services in the United States, the income elasticities generally fall in the range 0.3 to 0.6 (Cutler, Elmendorf and Zeckhauser, 1993). With charitable giving by individuals, the income elasticities generally fall in the range of 0.4 to 0.8. (Clotfelter, 1985). The income elasticities in the contingent valuation literature vary with

[22]"Warm glow" is simply a red herring. I have seen no empirical evidence that people get a warm glow from voting to raise their own taxes, whether in real life or in a contingent valuation study.
[23]Milgrom (1993) also asserts that using contingent valuation to measure altruistic preferences creates double counting. His analysis has three flaws. First, it depends on the particular specification of the utility function, as Johansson (1992) notes; if the argument of the utility function is another's consumption rather than his utility, there is no double counting. Second, it derives its force from the auxiliary assumption that the respondent *does not realize* that the other people for whom he cares will have to pay, too; this is not a problem in a referendum format. Third, in many contingent valuation studies the object of the altruism is often wildlife—sea otters, for example. Since those creatures are *not* surveyed, the issue of double counting is moot.

the item being valued, but are generally in this same range (Kristrom and Riera, 1994).

The term "embedding effect," introduced by Kahneman and Knetsch (1992), has come to mean several different things. The general notion is captured in the (mis)conception that, with contingent valuation, you get the same willingness-to-pay if you value one lake, two lakes, or ten lakes.[24] This combines three distinct notions. One assertion, which arises when the object of preference is thought to be simply the number of lakes, is that willingness-to-pay varies inadequately with changes in the scale or scope of the item being valued. This is a scope effect. Alternatively, if each lake is seen as a separate argument in the utility function, then the assertion is that a given lake has quite different value if it is first, second or tenth in a set of items to be valued—it gets a high value when the first, but it adds little or nothing to total value when second or tenth. This is a sequencing effect. Thirdly, with either preference structure, the willingness-to-pay for a composite change in a group of public goods may be less than the sum of the willingnesses-to-pay for the individual changes separately. This is a sub-additivity effect.

The question of how willingness-to-pay varies with the scale or scope of the item being valued in a contingent valuation survey has long been considered, starting with Cicchetti and Smith (1973) who elicited hiker's values for trips in a Montana wilderness area and found that the willingness-to-pay for trips where other hikers were encountered on two nights was 34 percent lower than the willingness-to-pay for trips with no encounters. Many other studies have since reported comparable findings using both internal (within-subject) and external (split-sample) scope tests, including meta-analyses by Walsh, Johnson and McKean (1992) covering over 100 contingent valuation studies of outdoor recreation, and Smith and Osborne (1994) on 10 contingent valuation studies of air quality. Carson (1994) reviews 27 papers with split-sample tests of scope and finds a statistically significant effect of scope on willingness-to-pay in 25 of them.

The two exceptions are Kahneman and Knetsch (1992) and Desvousges et al. (1992). Critics of contingent valuation rely heavily on these two studies when asserting the absence of scope effects in contingent valuation.[25] Some of the problems with these two studies have already been noted, including their failure to use a closed-ended voting format, the after-the-fact provision of information in Kahneman and Knetsch's "top-down" procedure, and the use of

[24]Though widely believed, this is a myth. It may be traced to Kahneman (1986), which is usually cited as showing that respondents were willing to pay the same amount to clean up fishing lakes in one region of Ontario as in all of Ontario. His data actually show a 50 percent difference. Moreover, the survey involved a brief telephone interview using an open-ended willingness-to-pay question. It provided no detail on how and when the cleanup would occur. Respondents may not have seen cleaning up *all* the lakes as something likely to happen soon.

[25]Also, in their contingent valuation survey, Diamond et al. (1993, pp. 45–46) mention that, using a Kruskal-Wallis test, they found no difference in willingness-to-pay for three wilderness areas ranging in size from 700,000 to 1.3 million acres. If they had run a simple regression of willingness-to-pay on acreage, they would have found a significant scope effect.

brief shopping mall intercepts by Desvousges et al.[26] The latter elicited people's willingness-to-pay for preventing the deaths of migratory waterfowl. Three separate versions of the questionnaire said that 2,000, 20,000, and 200,000 out of 85 million birds die each year from exposure to waste-oil holding ponds that could be sealed under a new program. Respondents were told that the deaths amounted to *much less than* 1 *percent* of the bird population, to *less than* 1 *percent*, and to *about* 2 *percent*. If respondents focused on the relative impact on the population, it is hard to believe that they would have perceived any real difference among these percentages. The results of the scope test depend crucially on how much one trims the data to remove what are clearly outliers. With a 10 percent trim, one obtains a highly significant scope effect.[27] At any rate, even if one regards these two studies as highly credible evidence that respondents were insensitive to scope, they certainly do not represent the majority finding in the contingent valuation literature regarding the variation of willingness-to-pay with scope.

How much should willingness-to-pay vary with scope? Diamond (1993) asserts that economic theory requires it to increase *more than proportionately* with the number of bird deaths. The variables in his model are the number of birds originally in the population, q_0, the number at risk of dying, q_R, and the number of those that are saved, q_S. Let $q_F \equiv q_0 - q_R + q_S$. Diamond assumes that people should care only about q_F, the ultimate number of birds, not how many were alive initially, at risk, or saved. He also assumes preferences are quasiconcave in q_0. The two assumptions together imply *quasiconvexity* in q_R, which is what makes the elasticity of willingness-to-pay with respect to q_R greater than unity. The conclusion depends critically on the assumption of perfect substitution between q_0, q_S and $-q_R$. When contingent valuation data disconfirm this, Diamond dismisses the method. Others might be more inclined to believe the data and drop the assumption.[28]

With regard to sequencing and sub-additivity effects, these effects are certainly present in contingent valuation responses, but one expects them to occur, and they can be explained in terms of substitution effects and diminish-

[26]Other questions about Kahneman and Knetsch are raised by Harrison (1992) and Smith (1992).

[27]How the survey was administered clearly affected the results. Schkade and Payne (1993) used the same questionnaire as Desvouges et al., but slowed respondents down and made them think about their answer. Their data show a different pattern of willingness-to-pay responses, and a significant relationship between willingness-to-pay and the percentage of birds killed (Haneman, 1994b).

[28]Some, while not sharing Diamond's extreme position on the elasticity of willingness-to-pay, still hold that contingent valuation responses vary inadequately with scale. People's perceptions undoubtedly differ from objective measures of attributes. But this is not just a feature of contingent valuation. In psychophysics, it has been known since the 1880s that there is a general tendency for judgments of magnitude to vary inadequately. Observers standing at a distance overestimate the height of short posts, and underestimate that of tall ones; people reaching quickly for an object overestimate small distances and angles, and underestimate large ones; subjects matching loudness of a tone to a duration overestimate the loudness of short tones, and underestimate the loudness of long ones; people overestimate infrequent causes of death, and underestimate frequent ones; small probabilities are overestimated, large ones underestimated (Poulton, 1989). This "response contraction bias" in judgment or rating is an authentic feature of how people perceive the world, not an artifact of contingent valuation.

ing marginal rates of substitution. When the quality of one lake improves, you value an improvement in a second lake *less* if the lakes are what Madden (1991) calls R-substitutes, and *more* if they are R-complements. Far from being inconsistent with economic preferences (Diamond et al., 1993, pp. 48–49), sub-additivity is likely to be the norm: while all goods cannot be R-complements, Madden shows they *can* all be R-substitutes.[29] Similarly, R-substitution explains sequence effects: if the lakes are R-substitutes, the willingness-to-pay for an improvement in one lake is *lower* when it comes at the end of a sequence of changes in lake improvements than at the beginning while the willingness-to-accept for the change in the lake is *higher* when it comes later in a sequence (Carson, Flores and Hanemann, 1992).[30] It should come as no surprise that the value of one commodity changes when the quantity of another varies: in other words, that willingness-to-pay depends on economic context.[31]

For many economists, the ultimate argument against contingent valuation is that it violates the habitual commitment of the profession to revealed preference. Three points should be noted. First, one must distinguish between private market goods and public goods. Revealed preference is harder to apply to the latter, especially when they are national rather than local public goods (Cropper, 1994). Second, revealed preference is not foolproof, either. It involves an extrapolation from observation of particular choices to general conclusions about preference. One relies on various auxiliary assumptions to rule out factors that might invalidate the extrapolation. Those assumptions are not themselves verifiable if one is restricted to observed behavior. This can sometimes make revealed preference a relatively hypothetical undertaking.[32]

[29]If the intention of the Diamond et al. (1993) contingent valuation survey was to test the adding-up of willingness-to-pay, it was strangely designed for the purpose. The survey stated that there were 57 federal wilderness areas in the Rocky Mountain states, without identifying them, and said that there now was a proposal to open these to commercial development. In one version, respondents were told that seven unidentified areas had already been earmarked for development, and were asked their willingness-to-pay to protect an eighth area, identified as the Selway Bitterroot Wilderness. In another, respondents were told that eight unnamed areas had been earmarked for development and asked their willingness-to-pay to protect a ninth area, identified as the Washakie Wilderness. In a third version, respondents were told that seven unnamed areas had been earmarked for development and asked their willingness-to-pay to protect two areas identified as Selway and Washakie. In all three cases, respondents were not told the identity or fate of the other 48 or 49 areas. Given that respondents were not indifferent among wilderness areas, as evidenced by the regression mentioned in note 25, I leave it to the reader to decide whether the surveys constitute a sensible basis for testing the adding up of willingness-to-pay.

[30]In natural resource damages, where willingness-to-accept is the relevant welfare measure, this implies that the usual practice of taking the injured resource as the first item in any possible valuation sequence is a conservative procedure.

[31]The practical implications are that, when one values a program, it be placed in whatever sequence applies under the circumstances, and that one take care when extrapolating results in a benefits transfer exercise because the values might change with the difference in circumstances (Hoehn and Randall, 1989).

[32]Revealed preference estimates are sensitive to the measurement of price, which is often uncertain and precarious for disaggregated commodities (Pratt, Wise and Zeckhauser, 1979; Randall, 1994). The price at which demand falls to zero, needed to estimate consumer's surplus, may lie outside the range of the observed data and be estimated inaccurately (for example, one knows travel cost only for participants, or one believes that participants and nonparticipants have different preferences). This can cause revealed preference to produce a less reliable estimate of use value than contingent

Third, there is no reason why observing people's behavior and asking them about behavioral intentions and motives should be mutually exclusive. Fathoming human behavior is never easy; one should utilize every possible source of information.

Above all, one should take a balanced view of the difficulties with each approach. As Sen (1973, p. 258) wrote, "we have been too prone, on the one hand, to overstate the difficulties of introspection and communication and, on the other, to underestimate the problems of studying preferences revealed by observed behavior." In the debate on contingent valuation, critics have shown a tendency to employ simplistic dichotomies. Surveys of attitudes are fallible and subject to the vagaries of context and interpretation; surveys of behavior are unerring. In the market place, people are well informed, deliberate, and rational. Outside it, they are ignorant, confused, and illogical. As consumers, people can be taken seriously; as voters, they cannot. In particular instances, these assertions may be correct. As generalizations, however, they are a caricature.

Conclusions

When cost-benefit analysis started in the United States in the 1930s, economic valuation was generally perceived in terms of market prices. To value something, one ascertained an appropriate market price, adjusted for market imperfections if necessary, and then used this to multiply some quantity. Two things changed this. The first was the recognition, prompted by the "new welfare economics" of the 1940s and especially Hotelling's paper on public utility pricing, that the appropriate welfare criterion is maximization of aggregate consumers' plus producers' surplus. While market prices can safely be used to value marginal changes for market commodities, the impact of nonmarginal changes is measured by the change in areas under demand and supply curves. The second development was Samuelson's theory of public goods and his finding that their valuation must be based on vertical aggregation of individual demand curves.

Together, these developments led to an important paradigm shift—one that contributed directly to the emergence of nonmarket valuation and is still evident in the current debate on contingent valuation.[33] This shift changed the focus of valuation away from market prices towards demand and supply functions as the underlying repositories of value. These functions are behavioral relations, and the implication of the paradigm shift was that economics is

valuation (Hanemann, Chapman and Kanninen, 1993). With other variables there may be inadequate variation in the data (for example, attributes are correlated across brands). Hence, revealed preference data alone may yield a less reliable estimate of demand functions than contingent valuation choice data, and one may need to combine both types of data for best results (Adamowicz, Louviere and Williams, 1994).

[33]For an account of the development of nonmarket valuation generally, see Hanemann (1992).

not just the study of markets, but more generally the study of human preferences and behavior.

The conceptual link to nonmarket valuation is the recognition that, while a demand curve is not observable if there is no market for a commodity, there still exists a latent demand curve that perhaps can be teased out through other means. Indirect methods are one approach to doing this, and contingent valuation is another. In both cases, the details of implementation have a large impact on the quality of the results.

Faced with the assertion that contingent valuation surveys can *never* be a reliable source of information either for benefit cost analysis or for damage assessment, the NOAA Panel rejected this as unwarranted. Two years later, there is now even more evidence from recent studies and literature analyses to support the Panel's conclusion. However, it would be misleading for me to suggest that contingent valuation surveys can be made to work well in all circumstances. I am sure situations could exist where a contingent valuation researcher might be unable to devise a plausible scenario for the item of interest. Nor would I wish to argue that all contingent valuation surveys are of high quality. The method, though simple in its directness, is in fact difficult to implement without falling into various types of design problems that require effort, skill and imagination to resolve. Each particular study needs to be scrutinized carefully. But the same is true of any empirical study.

While I believe in the feasibility of using contingent valuation to measure people's value for the environment, I do not mean to advocate a narrow benefit-cost analysis for all environmental policy decisions, nor to suggest that everything can or should be quantified. There will be cases where the information is inadequate, the uncertainties too great, or the consequences too profound or too complex to be reduced to a single number. I am well aware of the fallacy of misplaced precision. But this cuts both ways. It also applies to those who suggest that it is better not to measure nonuse values at all than to measure them through contingent valuation. I reply to such critics by quoting Douglass North: "The price you pay for precision is an inability to deal with real-world issues" (*Wall Street Journal*, 7/29/94).

Is expert judgment an alternative to contingent valuation? Experts clearly play the leading role in determining the physical injuries to the environment and in assessing the costs of clean-up and restoration. Assessing what things *are worth* is different. How the experts know the value that the public places on an uninjured environment, without resort to measurement involving some sort of survey, is unclear. When that public valuation is the object of measurement, a well-designed contingent valuation survey *is* one way of consulting the relevant experts—the public itself.

■ *I want to thank Richard Carson, Jon Krosnick, Robert Mitchell, Stanley Presser and Kerry Smith for their helpful comments, and Nicholas Flores and Sandra Hoffmann for excellent assistance. I also thank the editors, without whom this paper would be far longer.*

References

Adamowicz, W., J. Louviere, and M. Williams, "Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities," *Journal of Environmental Economics and Management*, 1994, *26*, 271–92.

Ajzen, Icek, and Martin Fishbein, *Understanding Attitudes and Predicting Social Behavior*. New Jersey: Prentice-Hall, Inc., 1980.

Anonymous, "'Ask a Silly Question...' Contingent Valuation of Natural Resource Damages," *Harvard Law Review*, June 1992, *105*, 1981–2000.

Arrow, Kenneth, J., *Social Choice and Individual Values*, 2nd ed., New Haven: Yale University Press, 1963.

Arrow, Kenneth, et al., *Report of the NOAA Panel on Contingent Valuation*, Washington, D.C.: January 1993, p. 41.

Atkinson, A. B., and J. Micklewright, "On the Reliability of Income Data in the Family Expenditure Survey, 1970–1977," *Journal of the Royal Statistical Society* (A), 1983, *146(I)*, 33–53.

Bate, Roger, "Pick a Number: A Critique of Contingent Valuation Methodology and Its Application in Public Policy." Competitive Enterprise Institute, Environmental Studies Program, Washington, D.C., January 1994.

Bateman, Ian, and Ken Willis (eds.), *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EC and Developing Countries*. Oxford, UK: Oxford University Press, forthcoming.

Becker, Gary S., "Nobel Lecture: The Economic Way of Looking at Behavior," *Journal of Political Economy*, June 1993, *101(3)*, 385–409.

Bettman, James R., "Processes of Adaptivity in Decision Making," *Advances in Consumer Research*, 1988, *15*, 1–4.

Bettman, J. R., and M. A. Zins, "Constructive Processes in Consumer Choice," *Journal of Consumer Research*, September 1977, *4*, 75–85.

Bishop, Richard C., and Thomas A. Heberlein, "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics*, December 1979, *61*, 926–30.

Bishop, Richard C., and Thomas A. Heberlein, "The Contingent Valuation Method." In Johnson, Rebecca L., and Gary V. Johnson, eds., *Economic Valuation of Natural Resources: Issues, Theory, and Applications*, Boulder: Westview Press, 1990, 81–104.

Bohm, Peter, "Estimating Demand for Public Goods: An Experiment," *European Economic Review*, 1972, *3*, 111–30.

Boyce, R. R., et al., "Experimental Evidence of Existence Value in Payment and Compensation Contexts." Paper presented at the USDA W-133 Annual Meeting, San Diego, California, February 1989.

Boyle, Kevin J., Michael P. Welsh, and Richard C. Bishop, "The Role of Question Order and Respondent Experience in Contingent-Valuation Studies," *Journal of Environmental Economics and Management*, 1993, *25*, S-80-S-99.

Bradburn, Norman M., Lance J. Rips, and Steven K. Shevell, "Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys," *Science*, April 1987, *236*, 157–161.

Briscoe, John, et al., "Toward Equitable and Sustainable Rural Water Supplies: A Contingent Valuation Study in Brazil," *World Bank Economic Review*, May 1990, *4*, 115–34.

Brookshire, David S., Mark A. Thayer, William D. Schulze, and Ralph C. d'Arge, "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches," *American Economic Review*, 1982, *72*, 165–77.

Carson, Richard T., "Contingent Valuation Surveys and Tests of Insensitivity to Scope." Paper presented at the International Conference on Determining the Value of Nonmarketed Goods: Economic Psychological, and Policy Relevant Aspects of Contingent Valuation Methods, Bad Hamburg, Germany, July 1994.

Carson, Richard T., and Nicholas E. Flores, "Another Look at 'Does Contingent Valuation Measure Preferences: Experimental Evidence'—How Compelling is the Evidence?" Economics Department, University of California, San Diego, December 1993.

Carson, R., N. Flores, and W. M. Hanemann, "On the Creation and Destruction of Public Goods: The Matter of Sequencing," working paper 690, Agricultural and Resource Economics, University of California, Berkeley, 1992.

Carson, Richard T., Nicholas E. Flores, Kerry Martin and Jennifer Wright, "Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods," Discussion Paper 94-07, University of California, San Diego, May 1994a.

Carson, Richard T., W. Michael Hanemann, and Robert Cameron Mitchell, "The Use of Simulated Political Markets to Value Public Goods," Economics Department, University of California, San Diego, October 1986.

Carson, Richard T., Kerry Martin, Jennifer Wright, "A Note on the Evidence of the Temporal Reliability of Contingent Valuation Estimates," working paper, University of California, San Diego, Economics Department, July 1994b.

Carson, Richard T., and Robert Cameron Mitchell, "The Value of Clean Water: The Public's Willingness to Pay for Boatable, Fishable, and Swimmable Quality Water," *Water Resources Research*, 1993, *29*, 2445–54.

Carson, R., et al., *A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill*, Report to the Attorney General of Alaska, Natural Resource Damage Assessment, Inc. La Jolla, CA, November 1992.

Carson, Richard T., et al., *A Bibliography of Contingent Valuation Studies and Papers*, Natural Resource Damage Assessment, Inc., La Jolla, CA, March 1994c.

Chase, S. B., ed., *Problems in Public Expenditure Analysis*. Washington, D.C.: Brookings Institution, 1968.

Cicchetti, Charles J., and V. Kerry Smith, "Congestion, Quality Deterioration, and Optimal Use: Wilderness Recreation in the Spanish Peaks Primitive Area," *Social Science Research*, 1973, *2*, 15–30.

Ciriacy-Wantrup, S. V., "Capital Returns from Soil-Conservation Practices," *Journal of Farm Economics*, November 1947, *29*, 1188–90.

Clotfelter, Charles T., *Federal Tax Policy and Charitable Giving*. Chicago: The University of Chicago Press, 1985.

Cronin, Thomas E., *Direct Democracy: The Politics of Initiative, Referendum, and Recall*. Cambridge: Harvard University Press, 1989.

Cropper, Maureen L., "Comments on Estimating the Demand for Public Goods: The Collective Choice and Contingent Valuation Approaches." Paper presented at the DOE/EPA Workshop on "Using Contingent Valuation to Measure Non-Market Values," Herndon, VA, May 19–20, 1994.

Cummings, Ronald G., "Relating Stated and Revealed Preferences: Challenges and Opportunities." Paper presented at the DOE/EPA Workshop on "Using Contingent Valuation to Measure Non-Market Values," Herndon, VA, May 19–20, 1994.

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

Cummings, Ronald G., Glenn W. Harrison, and E. E. Ruström, "Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible?" Economics Working Paper Series, B-92-12, Division of Research, College of Business Administration, The University of South Carolina, February 1993.

Cutler, David, Douglas W. Elmendorf, and Richard J. Zeckhauser, "Demographic Characteristics and the Public Bundle," National Bureau of Economic Research, Cambridge, NBER Working Paper No. 4283, February 1993.

Desvousges, William H., et al., *Measuring Nonuse Damages Using Contingent Valuation: An Experimental Evaluation of Accuracy*. North Carolina: Research Triangle Institute Monograph, 1992.

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

DiBona, Charles J., "Assessing Environmental Damage," *Issues in Science and Technology*, Fall 1992, *8*, 50–54.

Dickie, M. A. Fisher, and S. Gerking, "Market Transactions and Hypothetical Demand Data: A Comparative Study," *Journal of American Statistical Association*, March 1987, *82*, 69–75.

Duffield, John W., and David A. Patterson, "Field Testing Existence Values: An Instream Flow Trust Fund for Montana Rivers." Presented at the American Economics Association Annual Meeting, New Orleans, Louisiana, January 4, 1991.

Ferber, Robert, and Robert A. Piskie, "Subjective Probabilities and Buying Intentions," *Review of Economics and Statistics*, August 1965, *47*, 322–25.

Fiorina, Morris P., *Retrospective Voting in American National Elections*. New Haven: Yale University Press, 1981.

Fishkin, J. S., *Democracy and Deliberation: New Directions for Democratic Reform*. New Haven: Yale University Press, 1991.

Freeman, A. Myrick, *The Measurement of Environment and Resource Values: Theory and*

Method. Washington, D.C.: Resources for the Future, 1993.

Frisch, Deborah, "Reasons for Framing Effects," *Organizational Behavior and Human Decision Processes*, 1993, *54*, 399–429.

Groves, Robert M., *Survey Errors and Survey Costs*. New York: John Wiley and Sons, 1989.

Hanemann, W. Michael, "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses," *American Journal of Agricultural Economics*, August 1984, *66*, 332–41.

Hanemann, W. Michael, "Preface: Notes on the History of Environmental Valuation in the USA." In Navrud, Stale, ed., *Pricing the Environment: The European Experience*. Oxford, UK: Oxford University Press, 1992.

Hanemann, W. Michael, "Contingent Valuation and Economics," Working Paper No. 697, Giannini Foundation of Agricultural and Resource Economics, University of California, Berkeley, February 1994a. To appear in Willis, Ken, and John Corkindale, eds., *Environmental Valuation: Some New Perspectives*, Wallingford, Oxon, UK: CAB International, forthcoming.

Hanemann, W. Michael, "Strictly For the Birds: A Re-examination of the Exxon Tests of Scope in CV," working paper, Giannini Foundation of Agricultural and Resource Economics, University of California, Berkeley, August 1994b.

Hanemann, W. Michael, and B. J. Kanninen, "Statistical Analysis of CV Data." In Bateman, I., and K. Willis, eds. *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EC and Developing Countries*. Oxford: Oxford University Press, forthcoming.

Hanemann, W. Michael, David Chapman, and Barbara Kanninen, "Non-Market Valuation Using Contingent Behavior: Model Specification and Consistency Tests." Presented at the American Economic Association Annual Meeting, Anaheim, California, January 6, 1993.

Hanemann, W. M., J. Loomis, and B. Kanninen, "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation," *American Journal of Agricultural Economics*, 1991, *73*, 1255–63.

Harrison, Glenn W., "Valuing Public Goods with the Contingent Valuation Method: A Critique of Kahneman and Knetsch," *Journal of Environmental Economics and Management*, 1992, *23*, 248–57.

Hausman, J. A., ed., *Contingent Valuation: A Critical Assessment*. New York: North-Holland, 1993.

Hoehn, J. P., and A. Randall, "Too Many Proposals Pass the Benefit Cost Test," *American Economic Review*, June 1989, *79*, 544–51.

Infosino, William J., "Forecasting New Product Sales from Likelihood of Purchase Ratings," *Marketing Science*, Fall 1986, *5*, 372–384.

Johansson, Per-Olov, "Altruism in Cost-Benefit Analysis," *Environmental and Resource Economics*, 1992, *2*, 605–13.

Jones-Lee, Michael W., M. Hammerton, and P. R. Philips, "The Value of Safety: Results of a National Sample Survey," *Economic Journal*, March 1985, *95*, 49–72.

Juster, F. Thomas, *Anticipations and Purchases: An Analysis of Consumer Behavior*. Princeton: Princeton University Press, 1964.

Juster, F. Thomas, and Frank P. Stafford, "The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement," *Journal of Economic Literature*, 1991, *29*, 471–522.

Kahneman, Daniel, "Valuing Environmental Goods: An Assessment of the Contingent Valuation Method: The Review Panel Assessment." In Cummings, R. G., D. S. Brookshire, W. D. Schulze, et al., eds., *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*. Totowa, New Jersey: Rowman & Allanheld, 1986, 185–94.

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

Kahneman, Daniel, and Ilana Ritov, "Determinants of Stated Willingness to Pay for Public Goods: A Study in the Headline Method," unpublished, Department of Psychology, University of California, Berkeley, 1993.

Kanninen, B. J., "Optimal Experimental Design for Double-Bounded Dichotomous Choice Contingent Valuation," *Land Economics*, May 1993, *69*, 128–46.

Kelly, S., and T. W. Mirer, "The Simple Act of Voting," *American Political Science Review*, 1974, *68*, 572–91.

Kemp, Michael A., and Christopher Maxwell, "Exploring a Budget Context for Contingent Valuation Estimates." In Hausman, J. A., ed., *Contingent Valuation: A Critical Assessment*. New York: North-Holland, 1993, 217–69.

Knetsch, J. L., and R. K. Davis, "Comparisons of Methods for Recreation Evaluation." In Kneese A. V., and S. C. Smith, eds., *Water Research*, Baltimore: Resources for the Future Inc., Johns Hopkins Press, 1966, 125–42.

Kriström, Bengt, and Pere Riera, "Is the Income Elasticity of Environmental Improvements Less Than One?" Paper presented at the Second Conference on Environmental Economics, Ulvöng, Sweden, June 2–5, 1994.

Krosnick, Jon A., "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys," *Applied Cognitive Psychology*, 1991, *5*, 213–36.

Krosnick, Jon A., Fan Li, and Darrin R. Lehman, "Conversational Conventions, Order of Information Acquisition, and the Effect of Base Rates and Individuating Information on Social Judgments," *Journal of Personality and Social Psychology*, 1990, *59*, 1140–52.

Krutilla, John V., "Conservation Reconsidered," *American Economic Review*, September 1967, *57*, 777–86.

Lupia, Arthur, "Short Cuts versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections," working paper, Department of Political Science, University of California, San Diego, April 1993.

Madden, Paul, "A Generalization of Hicksian *q* Substitutes and Complements with Application to Demand Rationing," *Econometrica*, September 1991, *59*, 1497–1508.

Magleby, David B., *Direct Legislation, Voting on Ballot Propositions in the United States*. Baltimore and London: The John Hopkins University Press, 1984.

Maki, Atsushi, and Shigeru Nishiyama, "Consistency Between Macro- and Micro-Data Sets in the Japanese Household Sector," *Review of Income and Wealth*, 1993, *39*, 195–207.

Martin, Leonard L., and Abraham Tesser, eds., *The Construction of Social Judgments*. New Jersey: Lawrence Erlbaum Associates, chapter 2, 1992, 37–65.

Mathiowetz, Nancy A., and Greg J. Duncan, "Out of Work, Out of Mind: Response Errors in Retrospective Reports of Unemployment," *Journal of Business & Economic Statistics*, 1988, *6*, 221–29.

McCloskey, Donald, *The Rhetoric of Economics*. Madison: The University of Wisconsin Press, 1985.

McFadden, Daniel, and Gregory K. Leonard, "Issues in the Contingent Valuation of Environmental Goods: Methodologies for Data Collection and Analysis." In Hausman, J. A., ed., *Contingent Valuation: A Critical Assessment*. New York: North-Holland, 1993, 165–215.

McKelvey, Richard D., and Peter C. Ordeshook, "Information, Electoral Equilibria and the Democratic Ideal," *Journal of Politics*, 1986, *48*, 909–37.

Merton, Robert K., and Patricia L. Kendall, "The Focused Interview," *American Journal of Sociology*, 1946, *51*, 541–57.

Milgrom, Paul, "Is Sympathy an Economic Value? Philosophy, Economics, and the Contingent Valuation Method." In Hausman, J. A., ed., *Contingent Valuation: A Critical Assessment*. New York: North-Holland, 1993, 417–41.

Miller, Leslie A., and Theodore Downes-Le Guin, "Reducing Response Error in Consumers' Reports of Medical Expenses: Application of Cognitive Theory to the Consumer Expenditure Interview Survey," *Advances in Consumer Research*, 1990, *17*, 193–206.

Mitchell, Robert Cameron, and Richard T. Carson, *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, D. C.: Resources for the Future, 1989.

Navrud, Ståle, *Pricing the European Environment*. New York: Oxford University Press, 1992.

Neill, Helen R., et al., "Hypothetical Surveys and Real Economic Commitments," *Land Economics*, May 1994, *70*, 145–54.

Neisser, Urlic, *Cognitive Psychology*. Appleton-Century-Crofts, Educational Division, New York: Meredith Corporation, 1967.

Oates, W., "Comments on Estimating the Demand for Public Goods: The Collective Choice and Contingent Valuation Approaches." Paper presented at the DOE/EPA Workshop on Using Contingent Valuation to Measure Non-Market Values, Hemdon, VA, May 19–20, 1994.

Olshavsky, Richard W., and Donald H. Granbois, "Consumer Decision Making—Fact or Fiction?" *Journal of Consumer Research*, September 1979, *6*, 93–100.

Payne, J. W., J. R. Bettman, and E. J. Johnson, "Adaptive Strategy Selection in Decision Making," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1988, *14*, 534–52.

Poulton, E. C., *Bias in Quantifying Judgments*. Hove, UK: Lawrence Erlbaum Associates, 1989.

Pratt, John W., David A. Wise, and Richard Zeckhauser, "Price Differences in Almost

Competitive Markets," *Quarterly Journal of Economics*, May 1979, *93*, 189–212.

Randall, Alan, "A Difficulty with the Travel Cost Method," *Land Economics*, February 1994, *70*, 88–96.

Randall, Alan, Berry C. Ives, and Clyde Eastman, "Bidding Games for Valuation of Aesthetic Environmental Improvements," *Journal of Environmental Economics and Management*, 1974, *1*, 132–49.

Rose, Steven, *The Making of Memory: From Molecules to Mind*. New York: Anchor Books, Doubleday, 1992.

Rosen, S., "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, January-February 1974, *82*, 34–55.

Schelling, Thomas, "The Life You Save May Be Your Own." In Chase, S., ed., *Problems in Public Expenditure Analysis*. Washington, D.C.: Brookings Institution, 1968, 143–4.

Schkade, David A., and John W. Payne, "Where Do the Numbers Come From? How People Respond to Contingent Valuation Questions." In Hausman, J. A., ed., *Contingent Valuation: A Critical Assessment*. New York: North-Holland, 1993, 271–303.

Schuman, H., remarks in transcript of Public Meeting of the National Oceanic and Atmospheric Administration, Contingent Valuation Panel, Washington, D.C.: NOAA, Department of Commerce, August 12, 1992, p. 101.

Schuman, H., and S. Presser, *Questions and Answers in Attitude Surveys*. New York: Academic Press, 1981.

Seip, K., and J. Strand, "Willingness to Pay for Environment Goods in Norway: A Contingent Valuation Study with Real Payment," *Environmental and Resource Economics*, 1992, *2*, 91–106.

Sen, A. K., "Behavior and the Concept of Preference," *Economica*, August 1973, *40*, 241–59.

Sinden, J. A., "Empirical Tests of Hypothetical Biases in Consumers' Surplus Surveys," *Australian Journal of Agricultural Economics*, 1988, *32*, 98–112.

Smith, Tom W., "The Art of Asking Questions, 1936–1985," *Public Opinion Quarterly*, 1987, *51*, 21–36.

Smith, V. Kerry, "Arbitrary Values, Good Causes, and Premature Verdicts," *Journal of Environmental Economics and Management*, 1992, *22*, 71–89.

Smith, V. Kerry, "Lightning Rods, Dart Boards and Contingent Valuation," *Natural Resources Journal*, forthcoming 1994.

Smith, V. Kerry, and William H. Desvousges, *Measuring Water Quality Benefits*. Boston: Kluwer-Nijhoff Publishing, 1986.

Smith, V. Kerry, and Laura Osborne, "Do Contingent Valuation Estimates Pass a 'Scope' Test?: A Preliminary Meta Analysis." Presented at the American Economics Association Annual Meeting, Boston MA, January 5, 1994.

Sniderman, Paul M., "The New Look in Public Opinion Research." In Finifter, Ada W., ed., *Political Science: The State of the Discipline II*. Washington, D.C.: The American Political Science Association, 1993, 219–45.

Sniderman, Paul M., Richard A. Brody, and Phillip E. Tetlock, *Reasoning and Choice, Explorations in Political Psychology*. Cambridge: Cambridge University Press, 1991.

Tanur, Judith M., ed., *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*. New York: Russell Sage Foundation, 1992.

Theil, Henri, and Richard F. Kosobud, "How Informative Are Consumer Buying Intentions Surveys?" *Review of Economics and Statistics*, February 1968, *50*, 50–59.

Tversky, Amos, and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, 1974, *185*, 124–31.

Walsh, Richard G., Donn M. Johnson, and John R. McKean, "Benefits Transfer of Outdoor Recreation Demand Studies: 1968–1988," *Water Resources Research* 1992, *28*, 707–13.

Warshaw, Paul R., "Predicting Purchase and Other Behaviors from General and Contextually Specific Intentions," *Journal of Marketing Research*, February 1980, *17*, 26–33.

Wilson, Timothy D., and Sara D. Hodges, "Attitudes as Temporary Constructions." In Martin, L., and A. Tesser, eds., *The Construction of Social Judgments*. New Jersey: Lawrence Erlbaum Associates, chapter 2, 1992, 37–65.