WikiLeaks Document Release

Congressional Research Service

Report RL30242

NATURAL RESOURCES: ASSESSING NONMARKET VALUES THROUGH CONTINGENT VALUATION

Joseph Breedlove, Resources, Science, and Industry Division Intern

Updated June 21, 1999

Abstract. This report provides background on the nonmarket value of natural resources and the strengths and weaknesses of contingent valuation for estimating such values. Nonmarket values are increasingly being recognized as important in natural resource damage assessments and decisionmaking. This report describes contingent valuation, a survey technique often used to estimate nonmarket values, and examines its strengths and weaknesses.

WikiLeaks

CRS Report for Congress

Received through the CRS Web

Natural Resources:
Assessing Nonmarket Values
through Contingent Valuation

June 21, 1999

Joseph Breedlove
Graduate Student Intern
Under the Direction of Ross W. Gorte
Specialist in Natural Resources Policy
Resources, Science, and Industry Division

Congressional Research Service ❖ The Library of Congress

ABSTRACT

This report provides background on the nonmarket value of natural resources and the strengths and weaknesses of contingent valuation for estimating such values. Nonmarket values are increasingly being recognized as important in natural resource damage assessments and decisionmaking. This report describes contingent valuation, a survey technique often used to estimate nonmarket values, and examines its strengths and weaknesses. This report will not be updated.

Natural Resources: Assessing Nonmarket Values through Contingent Valuation

Summary

The role of nonmarket values in natural resource damage assessments and decisionmaking is being increasingly recognized. Numerous statutes direct federal agencies to provide goods and services efficiently, often necessitating a measure of nonmarket values. However, legislative issues have focused on damage assessment under Superfund, because the tax authorization under this law expired at the end of 1996, and thus Congress may debate its reauthorization.

Including nonmarket values in damage assessment and decisionmaking can be highly controversial. Proponents assert that excluding (not estimating) such values understates total values affected, often substantially, and biases decisions in favor of development. Critics counter that the measurement methodology is weak, and that such measures are not comparable to traditional measures of utilitarian values, because resource use generates economic and social benefits beyond those measured by price and volume (the traditional measures of utilitarian value). Thus, they argue that including nonmarket values can lead to arbitrary assessments of damage.

Contingent valuation is a survey technique that is purported to estimate the nonmarket value of the specified goods and services. The regulations for cost and damage recovery under the federal Superfund program explicitly recognize the use of contingent valuation as a tool for estimating such values. Contingent valuation surveys have been used in numerous settings; three significant examples include: the valuation of air quality improvement at the Grand Canyon; the valuation of damages resulting from the *Exxon Valdez* oil spill in Alaska; and the valuation of benefits from altering Glen Canyon dam operations.

Contingent valuation surveys use hypothetical markets to replicate actual markets or referenda for respondents to reveal their preferences for a good. Typically, respondents are asked how much they would be willing to pay (in higher prices or in taxes) for a particular action. The results of such surveys can always be questioned, because of the array of possible measurement errors and biases, because of empirical evidence challenging their reliability and validity, and because of incompatibility with market-based use values. Nonetheless, nonuse values are real, and ignoring them could significantly understate total losses, since nonuse values can at times be substantial. Thus, using contingent valuation (and other methods) to estimate nonuse values is likely to continue and to become more controversial.

Contents

Natural Resources:
Assessing Nonmarket Values
through Contingent Valuation

Introduction

Contingent valuation is a survey method used to estimate the nonuse value of public goods — generally defined as the value people place on certain goods simply because those goods exist. Contingent valuation is becoming more widely used in appraising natural resource damages and in decisionmaking. Critics object to its use, arguing that such surveys of existence values are not comparable to the traditional measures of utilitarian values, because resource use generates economic and social benefits beyond those measured by price and quantity (the traditional measures of utilitarian value). This issue is of interest to Congress, because the regulations to implement Superfund allow the use of contingent valuation for measuring nonuse values in damage assessment. The Superfund tax has recently expired and its re-authorization may be debated in the 106th Congress. Other laws allow for nonuse values in damage appraisals and cost recovery, and federal resource management agencies are often required to balance values provided. Ultimately, Congress may be asked to decide whether and how nonuse values are to be included and balanced with use values in damage assessments and in resource management.

This report describes the use of contingent valuation surveys for estimating nonuse values. It contains a brief legislative background on the use of contingent valuation surveys under Superfund and other statutes providing for cost recovery for resource damages. That is followed by an overview of the economic theory behind measuring nonuse values, and then a brief discussion of three cases where contingent valuation surveys were used. The remainder of the report describes the contingent valuation method — survey design, reliability and measurement error, validity and bias, and three empirical critiques.

Legislative Background

Federal Resource Decisionmaking

As described in the following section, governments that own natural resources manage them for a number of purposes. Their actions are sometimes justified by the real or perceived failures or limitations in the markets for many of the goods and services provided by lands and resources. For example, the Reclamation Act of 1902 authorized the Secretary of the Interior to provide water to irrigate agricultural lands in the arid west when private developers were unable or unwilling to finance water

development. The purposes for building and operating large water projects have been expanded over the years to include municipal and other water supplies, down-stream flood control, recreational use, and fish and wildlife habitat. The Bureau of Reclamation and the Army Corps of Engineers, which constructs and operates dams mainly in the east and midwest, are required to allocate costs among the various beneficiaries, and thus implicitly to value the goods and services that the projects provide and that are not sold in markets.[1]

Similarly, management decisions for federal lands often address goods and services that are not sold in markets. The national forests managed by the Forest Service and the public lands managed by the Bureau of Land Management are to be administered for sustained yields of multiple uses: high regular or periodic outputs ("for outdoor recreation, range, timber, watershed, and wildlife and fish purposes") while maintaining the productivity of the land. The Federal Land Policy and Management Act of 1976 also specifies that the federal government should "receive fair market value of the use of the public lands and their resources unless otherwise provided for by statute." Because the level and mix of uses are mainly determined in public planning and by the annual congressional budget process, some analysts believe comparable valuation measures for the marketed commodities and the non-marketed goods and services could prove useful for approximating a socially efficient mix of outputs; this could include more astute use of markets and market signals, as well as nonmarket valuation techniques.

To date, contingent valuation has apparently not been used by the agencies for valuing the unmarketed goods and services in either of these contexts. The statutes and regulations governing these decisionmaking processes do not specify how to value the unmarketed goods and services, or to balance these values with marketed goods and services. However, the necessary allocations — explicit for water projects and implicit for land management — require some comparable valuation. Contingent valuation is a technique that might prove useful in some situations

Damage Assessment

In contrast to federal resource decisionmaking, federal damage assessment laws and regulations have been more explicit, with respect to assessing nonuse values that have been damaged. Legislative issues have focused on resource damage assessment under Superfund — the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) of 1980[2] — authorizing federal cleanup of waste sites and recovery of damages. Damage measurement was delegated to the Department of the Interior, with little guidance in the act. The regulations allow some damages to be

---

[1] See, for example, the Water Resources Development Act of 1986 (WRDA; P.L. 99-662), which altered cost-sharing formulas for many Corps projects. Typically, authorizing legislation for water projects specify which costs are to be at least partly reimbursed by users (such as for irrigation and for municipal and industrial use) and which are to be borne by the federal government (such as for recreation use and fish and wildlife habitat); these cost shares are allocated based on the relative benefits produced. For more, see CRS Report 98-980 ENR, *Federal Sales of Natural Resources: Pricing and Allocation Mechanisms*.

[2] P.L. 96-510, 94 Stat. 2767, as amended. 42 U.S.C. 9601 *et seq*.

determined using "simplified assessments requiring minimal field observation." More complicated damages are determined using "alternative methodologies for conducting assessments in individual cases to determine the type and extent of short and long-term injury and damages." The regulations allow measuring option and existence values only if no use values can be determined, and contingent valuation can only be used in such circumstances.[3] (The distinction between use and existence value is discussed below.) This effectively created a hierarchy, in which use values and market methods were preferred to nonuse values and nonmarket methods.[4]

Critics argue that these regulations underestimate damages, and the regulations have been challenged. In *State of Ohio v. U.S. Department of the Interior*, 10 states (including Ohio) and 3 environmental organizations challenged the regulations, claiming that resource damages would be underestimated using those procedures.[5] The court found that

> regulation prescribing hierarchy of methodologies by which lost use value of natural resources could be measured, that focuses exclusively on market values for such resources when market values are available, was not reasonable inter-pretation of CERCLA.

A utility, a manufacturing company, and a chemical trade organization also challenged the regulations, arguing that contingent valuation could not be labeled a "best available procedure," as required by §301(c)(2) of CERCLA. The court upheld this part of the regulations, stating that

> Department of the Interior's inclusion of contingent valuation as method-ology to be employed in assessing damages resulting from harm to natural resources ... was proper; contingent valuation process includes techniques of setting up hypothetical markets to elicit individual's economic valuation of natural resource, and the methodology qualified as best available procedure for determining damages flowing from destruction of or injury to natural resources if properly applied and structured to eliminate undue upward biases.

Other federal laws also provide for damage recovery, and thus may implicitly authorize contingent valuation for some values.[6] The Clean Water Act authorizes the government to act as a natural resources trustee to recover damages (originally equal to restoration or replacement costs) or hazardous discharges into navigable waters or

---

[3] 43 *C.F.R.* Part 11.

[4] Robert F. Copple, "NOAA's Latest Attempt at Natural Resource Damage Regulation: Simpler ... But Better?" *Environmental Law Reporter: News & Analysis*, v. 25, no. 12 (Dec. 1995): 10671-10677; and Erik D. Olson, "Natural Resource Damages in the Wake of the Ohio and Colorado Decisions: Where Do We Go From Here?" *Environmental Law Re-porter: News & Analysis*, v. 19, no. 12 (Dec. 1989): 10551-10557.

[5] *State of Ohio v. United States Department of the Interior*, 880 F.2d 432 (D.C. Cir. 1989).

[6] Frank B. Cross, "Natural Resource Damage Valuation," *Vanderbilt Law Review*, v. 42, no. 2 (March 1989): 269-341.

near the coastline.[7] Other federal legislation — such as the Deepwater Port Act of 1974, the Trans-Alaska Pipeline Act, and the Outer Continental Shelf Lands Act — also provide for recovery of damages, but also typically fail to specify which methods can and should be used for calculating damages.[8] Some state laws also allow damage recovery and provide various types and levels of coverage.

In 1990, Congress enacted the Oil Pollution Act, delegating natural resource damage assessment for oil discharges into navigable waters, adjoining shorelines, or the Exclusive Economic Zone to the Department of Commerce, National Oceanic and Atmospheric Administration (NOAA).[9] To assist in developing the regulations, NOAA commissioned a panel of economic experts, co-chaired by Nobel laureate economists Kenneth Arrow and Robert Solow, to evaluate the use of contingent valuation in determining nonuse values for natural resource damage assessment. The panel concluded that contingent valuation could be used for such a purpose, subject to numerous conditions; the final report was published in the *Federal Register* on January 15, 1993.[10]

Economic Theory of Measuring Values

The capitalist economic system of the United States generally relies on transactions between producers and consumers in free markets to determine the outputs of goods and services. Prices established within this private exchange system are the basis for allocating land, labor, and capital among producers, and goods and services among consumers; prices are also the standard measure of value for such privately traded goods and services.

Governments often intervene in private transactions — by regulating private actions, by altering incentives, or by owning factors of production — to alter market results that are deemed socially or politically unacceptable. Reasons cited for government intervention include two classical market limitations (also called market failures): externalities and public goods. *Externalities* occur when private transactions between producers and consumers affect third parties (those not involved in the exchange), and those effects are not taken into account in the exchange.[11] For example, timber sales are exchanges between landowners and timber processors, but

---

[7] P.L. 92-500, 86 Stat. 816, as amended. 33 U.S.C. 1321(f).

[8] Respectively: P.L. 93-627, 88 Stat. 2126 (33 U.S.C. 1501, *et seq.*); P.L. 93-153, 87 Stat. 576 (43 U.S.C. 1651, *et seq.*); and P.L. 95-372, 92 Stat. 629 (43 U.S.C. 1301, *et seq.*).

[9] P.L. 101-380, 104 Stat. 484. 33 U.S.C. 2701, *et seq.*

[10] U.S. Dept. of Commerce, National Oceanic and Atmospheric Admin., "Natural Resource Damage Assessments Under the Oil Pollution Act of 1990," *Federal Register*, v. 58, no. 10 (Jan. 15, 1993): 4601-4614. (Hereafter referred to as the NOAA Panel.)

[11] Some mechanisms, such as litigation, can force producers and consumers to consider third-party effects. Markets respond to such mechanisms by internalizing at least some of the costs; for example, safety devices to protect consumers from unsafe automobiles have at least partly resulted from insurance litigation (as well as from government regulation), and the cost of these devices are now internal to the private transaction between producers and buyers. When third-party costs are fully internalized, no externalities occur.

sales can affect other people by altering animal habitats, the quantity and quality of water flows, and other land and resource conditions. Externalities are considered market limitations, because the exchanges ignore some of the costs (or benefits) they impose on society, and thus may result in more (or less) production than is socially desirable.

The second classical market problem is *public goods*: goods and services which can be used by one person without affecting the amount available for others and which are provided to all, because individuals cannot be excluded from the benefits. (Such goods are also called nondepletable or nonrival goods.) If the good is provided at all, it is not possible to exclude anyone from obtaining the benefits; for example, if we have national defense, all citizens have the same amount available to them, no matter how great a demand they have for it or how much they pay. Two additional aspects are common (but not required) in public goods: indivisibility and high transaction costs. Indivisibility results when the good or service cannot be divided among users; for example, individually owned pieces of the Statue of Liberty would not make much sense, because its value is in its entirety, not in its constituent pieces. High transaction costs occur when the owner or producer has difficulty controlling (and therefore charging for) the use or enjoyment; for example, the sheer size and accessibility of many lakes would prevent effective private control and access fees. In the most extreme forms, individuals *cannot* be excluded from receiving benefits. Private transactions in public goods may result in market "failures," because the possibility of simultaneous use, the indivisibility, and the difficulty or impossibility of controlling benefits make profitable private exchange ineffective, and thus, fewer public goods would be provided by private markets than are considered socially desirable.

Nonuse Values

As discussed above, public goods (*e.g.*, the Statue of Liberty) provided by the government are often used (*e.g.*, for recreation). Many public goods also have nonuse values — the value individuals place on goods or services, which they do not consume directly, often because an amenity or resource simply exists (known as existence value). It appears to have been first described by John Krutilla in 1967.[12] Evidence of such values is illustrated by voluntary contributions to a multitude of efforts and organizations; people are often willing to contribute time and money for things they feel have social value (*e.g.*, public television). For individuals, existence value can be inherent (valued solely because it exists) or for the future (valued because future generations will have it available, even if it is never used; this is sometimes known as bequest value).[13]

When uncertainty is considered, another form of existence value may result. People may be willing to pay for the option of using a good or service in the future,

---

[12] John V. Krutilla, "Conservation Reconsidered," *American Economic Review*, v. 56 (1967): 777-786.

[13] Robert Cameron Mitchell and Richard T. Carson, *Using Surveys To Value Public Goods: The Contingent Valuation Method* (Washington, DC: Resources for the Future, 1989), p. 62. (Hereafter referred to as Mitchell and Carson.)

typically at a particular price; for example, although you may not currently want to visit a recreational site, you may be willing to pay to have the option to visit it later. (This is often known as option value.) People may also be willing to pay for delays until better information is available; for example, an individual may be willing to pay to delay a project that may cause some irreversible effects that are not fully understood, if further information is likely to eliminate or to clarify those effects.[14] On the other hand, markets exist for option values of depletable goods and services, and are generally considered an efficient means of valuing future options and compensating the owner for maintaining that option.

Measuring Value

In a market economy, private goods and services are exchanged between producers and consumers. The standard measure of value for such goods and services is the price at which the exchange occurs willingly. Many government-provided goods, however, cannot be valued so readily, because they are not traded in markets and thus have no market price as a sign of the willingness of users to pay for the good.[15] Nonetheless, having values for such government-provided goods to approximate market prices can be useful for comparing alternative government programs that provide goods, both to improve government efficiency and to achieve a balance when the production of various goods conflict (*e.g.*, in situations where timber pro-duction compromises production of clean water from federal lands).

When market prices do not exist for government-provided goods, alternative methods must be used if one is to estimate the benefits of the public project, and to encompass total value: both use and nonuse values. One common approach — physical linkages — uses damage functions to estimate changes in direct use values, but it does not yield a complete measure of benefits, because it does not measure existence values. The second general approach — behavioral linkages — aggregates individual be-haviors to estimate values. Several approaches are possible, based on whether some relevant market behavior can be observed (observed *v.* hypothetical markets) and on how individual preferences for the good in question can be revealed (direct *v.* indirect measures). Table 1 shows these possibilities.

Observed/direct methods examine market behavior to estimate the value of a particular good directly. Referenda can measure popular support for a particular public project using a voting format. Simulated markets can determine a market price for a good by setting up an experimental market, such as a simulated market for hunting permits. If parallel private markets exist (*e.g.*, hunting clubs with exclusive hunting rights on certain lands), they can be used to estimate the value of a government-provided good. These methods are most effective when the value of the good

---

[14] Mitchell and Carson, pp. 69-74.

[15] This is not to say that markets cannot exist for many government-provided goods, but rather that society has chosen not to use markets for allocating those resources and for determining efficient production levels. Livestock grazing on federal lands, for example, is not allocated among ranchers by market decision, and is priced at an administratively determined fee, not at a rate set by a market for such grazing.

Table 1. Behavior-Based Methods of Valuing Unmarketed Goods

	Direct Measures	Indirect Measures
Observed behavior	Referenda Simulated markets Parallel private markets	Household production Hedonic pricing Actions of bureaucrats or politicians
Hypothetical markets	Contingent valuation Allocation game with tax refund Spend more-same-less survey question	Contingent ranking Willingness-to-pay (or -to-accept-compensation) Allocation games Priority evaluation technique Conjoint analysis Indifference curve mapping

Source: Mitchell and Carson, p. 75.

in question is primarily derived from its use (including the option for future use), but may be inadequate for goods with substantial nonuse values.

Observed/indirect methods examine market behavior for other, related goods to infer the value of the government-provided good. One commonly used household production function is the travel-cost method, which measures recreational benefits by calculating how much people are apparently willing to pay to visit a site (based on how far they travel to the site). Another technique is hedonic pricing, which includes property value and wage studies, to measure the value of certain character-istics of a location or a job, such as environmental amenities or health risks, which are capitalized in the property value or wage respectively. As with observed/direct methods, these methods may be inadequate for goods with substantial nonuse values. In addition, the severe data requirements and complex methodological considerations make implementation difficult.

Hypothetical/direct methods attempt to directly estimate the benefit of a good or service using a hypothetical situation. Contingent valuation is one such technique. It surveys the affected population to elicit the willingness-to-pay (or willingness-to-accept-compensation) for a change in the amount of a good provided (or available). Another technique is to survey citizen opinions concerning the adequacy of current spending levels for public projects. Such methods are useful, because they measure both use and existence values for public goods, and they can support the estimation of demand curves, and thus be more comparable to market prices for private goods. However, the rigorous data and methodological requirements make these techniques difficult to use fairly.

Hypothetical/indirect methods rely on hypothetical markets to indirectly obtain values for the public good in question. Examples include contingent ranking or the hypothetical travel-cost method. Contingent ranking generates a list of sites in order of preference, which are then translated into values. Some researchers claim that contingent ranking is easier for respondents than attempting to assign values to

commodities, but the technique yields results that are less readily comparable to market prices for private goods and services.

Property Rights and Value

Private markets work because the owners of the various goods have the right to deny use of the good to those who don't pay. (Thus, private goods are also called excludable goods.) In contrast, public goods are characterized partly by the inability of owners to exclude beneficiaries who do not pay. The rights of individuals to use a particular public good (or to have it exist) are often ill-defined. This lack of clarity leads to two different questions that can be asked to elicit the value of a particular public good: how much would you be willing to pay to acquire the proposed change (willingness-to-pay); or how much would you be willing to accept for the loss (of use, of quality, *etc*.) associated with the proposed change (willingness-to-accept-compensation). Willingness-to-pay essentially presumes that respondents do not have rights to the good in question, and must buy it. Willingness-to-accept presumes that respondents do hold property rights to the good in question, and that the right to change current conditions must be bought from them.[16]

Researchers have found differences between willingness-to-pay and willingness-to-accept amounts for the same public good, with the willingness-to-accept amount sometimes being substantially larger. Willingness-to-accept is likely to be higher for those goods that do not have close substitutes, and/or where people refuse to sell the good in question or want very high compensation for it because of some personal attachment. Furthermore, uncertainty and aversion to risk may lower responses of willingness-to-pay. Other theoretical explanations exist, such as that losses in utility (well-being) are valued differently than equivalent gains in utility.[17] The choice between willingness-to-pay and willingness-to-accept depends essentially on who has the right to the good in question. Since those rights are often not clearly established statutorily, the choice may not be obvious or indisputable.

Applications of Contingent Valuation

Contingent valuation is a survey technique to estimate nonuse values by asking respondents how much they are willing to pay or to accept for a change in the good in a hypothetical market framework. The first identified description of contingent valuation was in 1947 in an article by S. V. Ciriacy-Wantrup about measuring the benefits of preventing soil erosion.[18] Its first use was apparently by Robert Davis in his Ph.D dissertation in 1963, to measure the value of a recreation area to hunters and

---

[16] Daniel S. Levy, James K. Hammitt, Naihua Duan, Theo Downes-LeGuin, and David Friedman, *Conceptual and Statistical Issues in Contingent Valuation: Estimating the Value of Altered Visibility in the Grand Canyon*, MR-344-RC (Santa Monica, CA: Rand, 1995). (Hereafter cited as RAND Study.)

[17] Mitchell and Carson, pp. 30-41.

[18] Paul R. Portney, "The Contingent Valuation Debate: Why Economists Should Care," *Journal of Economic Perspectives*, v. 8, no. 4 (Fall 1994): 3-17.

wilderness advocates.[19] Since then, many studies have been conducted on a wide range of commodities, including environmental amenities and natural resources. Three significant studies are summarized here to illustrate several of the possible categories of nonuse values: altering visibility at the Grand Canyon; the *Exxon Valdez* oil spill; and changing the operating system at Glen Canyon Dam.

To improve visibility at the Grand Canyon, the U.S. Environmental Protection Agency (EPA) in 1991 required the Navajo Generating Station in Page, Arizona, to install scrubbers to reduce haze caused by sulfur dioxide emissions, at a cost of about $100 million annually. EPA was required to value the benefits, and chose to use a contingent valuation survey that estimated total annual benefits of $130-$250 mil-lion. The utility company responded with its own contingent valuation survey that estimated a benefit of only $50 million. Because of a long-standing interest in valuing public goods, RAND evaluated both studies and found several problems.[20] First, RAND asserted that willingness-to-accept should have been used (rather than willingness-to-pay), because the United States public allegedly holds property rights to Grand Canyon visibility. Second, RAND stated that the EPA study had incorrectly used visibility changes that were much larger than the actual changes. Finally, neither study was found to have used a large enough sample of respondents to be representative of the population of the United States.

Another significant use of contingent valuation was to measure lost existence values caused by the *Exxon Valdez* oil spill in 1989. Eleven million gallons of oil were spilled into Alaska's Prince William Sound, damaging surface waters, coastal land (including beaches and wetlands), marine plants, birds, fish, and marine mam-mals. A study for the State of Alaska[21] used willingness-to-pay; the authors argued that this was because of concerns about respondent beliefs about their rights, and because willingness-to-pay yields conservative estimates (compared to willingness-to-accept). A hypothetical scenario — another spill occurring again within the next 10 years if nothing was done to prevent it — asked how much respondents would be willing to pay to prevent similar damages. The result was a median willingness-to-pay of $31 per household, resulting in total damages of $2.8 billion ($31 each for an adjusted number of U.S. households), and was asserted to represent the lower bound of damages because conservative procedures were followed. An alternative study calculated recreation losses using the travel cost method.[22] It estimated a loss of $3.8 million for 1989, with no losses in 1990 and beyond. The authors of this second study expressed skepticism about contingent valuation results showing losses of several billion dollars, given such low values for recreation use losses.

---

[19] Portney, p 4.

[20] RAND Study.

[21] Richard T. Carson, Robert C. Mitchell, W. Michael Hanemann, Raymond J. Kopp, Stanley Presser, and Paul A. Ruud, *A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill*, A Report to the Attorney General of the State of Alaska (Nov. 10, 1992).

[22] Jerry A. Hausman, Gregory K. Leonard, and Daniel McFadden, "Assessing Use Value Losses Caused by Natural Resource Injury," in *Contingent Valuation A Critical Assessment*, ed. James A. Hausman (Amsterdam: North-Holland, 1993), pp. 341-359.

The third example is the reoperation of Glen Canyon Dam on the Colorado River in Arizona. Environmental impact statements were ordered from the Bureau of Reclamation by the Secretary of the Interior in 1989, to determine methods for operating the dam to protect downstream resources and Native American resources, on the belief that then-current operations were damaging those values. Altering operations to better approximate natural ecological conditions below the dam would, however, reduce power production, which was the principal purpose of the dam. Different operations — including varying the maximum and minimum flows, the variation in daily flow, and the rate of variation — were considered for improving resource conditions.[23] The Bureau contracted for a contingent valuation study to estimate the benefits of protecting downstream resources.[24] A mail survey was used to compare moderate flow fluctuation, low flow fluctuation, and seasonally adjusted steady flow with the baseline (no change). For the national sample, the aggregate annual values were $2.3 billion for the moderate flow fluctuation, $3.4 billion for the low flow fluctuation, and $3.4 billion for the seasonally-adjusted steady flow. For power users surveyed, the values were substantially lower: $62 million for the moderate flow fluctuation, $61 million for the low flow fluctuation, and $81 million for the seasonally adjusted steady flow.

The U. S. General Accounting Office (GAO) evaluated the contingent valuation study.[25] GAO reported that the recommendations from the NOAA panel of experts and Dillman's mail survey procedures (a standard for mail surveys) were generally followed, with two major exceptions: (1) in-person surveys are significantly better than the mail surveys used, and (2) the survey was six to eight pages longer than recommended. An unpublished rebuttal to the study by C.V. Jones (Economic Data Resources, Boulder, CO) and Mark Graham (Tri-State Generation and Transmission Association, Denver, CO) also had numerous criticisms, including that the national sample was not representative of the U.S. population and that the function used to estimate mean willingness-to-pay did not match the sample responses.[26] Others have suggested that the survey questions implied a significantly greater change in downstream environmental quality than was likely to result, at least for the first several years.[27]

---

[23] U.S. Dept. of the Interior, Bureau of Reclamation, Colorado River Studies Office, *Operation of Glen Canyon Dam: Draft Environmental Impact Statement. Summary* (Salt Lake City, UT: 1993), 65 pp.

[24] M.P. Welsh, R.C. Bishop, M.L. Phillips, and R.M. Baumgartner, *GCES Non-Use Value Study Final Report* (Madison, WI: Hagler Bailly Consulting. Sept. 8, 1995).

[25] U.S. General Accounting Office, *Bureau of Reclamation: An Assessment of the Environmental Impact Statement on the Operations of the Glen Canyon Dam*, GAO/RCED-97-12 (Washington, DC: U.S. Govt. Print. Off., Oct. 1996), 213 pp.

[26] C.V. Jones and Mark Graham, *Rebuttal to the GCES Non-Use Value Study Final Report,* unpublished report (June 4, 1996).

[27] Personal communication with John E. Schefter, Chief, Office of External Research, Water Resources Division, U.S. Geological Survey, Dept. of the Interior, on May 21, 1997.

The Contingent Valuation Method

Survey Design

A contingent valuation survey usually includes several parts: (1) an indication of property rights; (2) an emphasis on disposable income; (3) a description of the good to be valued; (4) the anticipated effects on the prices of other goods; (5) the payment mechanism; (6) the questions; and (7) data about the respondent.[28] But contingent valuation surveys differ from conventional surveys in several important ways. First, contingent valuation surveys usually value goods with which respond-ents have little experience. Second, contingent valuation surveys use hypothetical markets that must be believed and understood by respondents. Third, extra effort is required by respondents to determine which goods they prefer and how much they would pay to obtain them. These are some of the tests researchers face in designing valid and reliable contingent valuation surveys.

The NOAA panel of scientific experts[29] recommended in-person interviews over telephone surveys, which were, in turn, preferred to mail surveys. In-person surveys are more expensive than the other forms, but more complicated scenarios can be explained better using visual aids under this format. In contrast, telephone surveys are relatively inexpensive, but it may be difficult to explain the scenario in detail because phone calls are typically time-constrained. Mail surveys avoid interviewer bias, but are not subject to the same control that an in-person survey would generate.

The survey can simulate a private goods market or a political goods market. In a private market, people choose to buy varying amounts of the good at "market" prices. The average consumer is defined as the consumer who purchases the *mean* quantity of the good. In a political goods market, people vote as in a referendum on a public project, with payment coming through increased taxes. The average voter is the one who votes for the *median* quantity of the good. Potential problems exist for both formats. In the private goods model, a small number of individuals with high valuations can influence decisions and make everyone pay for the public good (suggesting that the *mode* might be preferable to the *mean* or *median*).[30] However, in a political goods market a majority can influence the decision to provide the good and not bear its full costs. The NOAA panel advocated the political goods market model, because it more closely resembles the way people already make decisions about government-provided goods.

---

[28] Mitchell and Carson, pp. 50-52.

[29] The panel was established by NOAA under the Oil Pollution Act of 1990 (P.L. 101-380); their report was printed in the *Federal Register*, v. 58, no. 10 (Jan. 15, 1993): 4601-4614.

[30] The *mean* is the average response, with the total (quantity or value) divided by the number of respondents. The *median* is the middle amount, with an equal number of higher and lower responses. The *mode* is the most common response (*i.e.*, the response given by the largest number of respondents).

Table 2 shows the choice of elicitation mechanisms (methods of asking about values) that are available. The methods can be separated depending on whether one or several questions are asked, and on whether the actual maximum willingness-to-pay (WTP) amount or a discrete approximation (a yes-or-no from each respondent to a specific amount) is received.

Table 2. Elicitation Mechanisms Available for Valuing Public Goods

	Actual WTP amount	Discrete indicator of WTP
Single question	Direct/open-ended question Payment card Sealed bid auction	Take-it-or-leave-it offer Spending question offer Interval checklist
Iterated series of questions	Bidding game Oral auction	Take-it-or-leave-it offer (with follow-up)

Source: Mitchell and Carson, p. 98.

Open-ended/direct questions do not give respondents an amount to consider; rather, they must come up with an amount on their own.[31] The payment card asks respondents to circle the value or specify a value that represents their maximum willingness-to-pay for the good, given a range of numbers listed on the survey form. It is not subject to starting-point biases from which some auction techniques suffer, but the range of numbers may bias responses.

Bidding games and auctions can also be used. In English auctions, bids are increased until the highest valuation is reached. In Dutch auctions, the initial price is set high and lowered until someone chooses to buy at that price.[32] Although these methods allow bidders to more carefully consider different prices, they may be subject to starting-point biases and may be expensive and time-consuming to implement.

The take-it-or-leave-it offer asks respondents to answer whether they would be willing to pay a set amount for a good. Although this format is likely to provide a more truthful valuation, it requires substantially more data than do other methods. Follow-up questions can improve the efficiency of this method, but they add their own statistical complications and potential biases.[33]

Additional criteria were suggested by the NOAA panel to cover other aspects of the survey instrument, including

- the willingness-to-pay question should correspond to a potential future event, not one that has already occurred;

---

[31] W. Michael Hanemann, "Valuing the Environment Through Contingent Valuation," *Journal of Economic Perspectives*, v. 8, no 4 (1994): 19-43.

[32] R.G. Cummings, D.S Brookshire, and W.D. Schulze, *Valuing Public Goods: An Assessment of the Contingent Valuation Method* (Totowa, NJ: Rowman and Allanhead Publishers, 1986), p. 39.

[33] Mitchell and Carson, pp. 99-104.

- the implications of any decision should be described in detail;
- respondents should be reminded of their budget constraints (that spending for a good means a reduction in other kinds of goods that can be purchased);
- respondents should also be made aware of possible substitutes for the good they are valuing; and
- follow-up questions should test how well respondents understood the scenario and to try to determine the motivation behind the different responses.[34]

The goal of contingent valuation researchers is to elicit responses that are both reliable and valid. A variety of techniques, such as pretesting tools and training interviewers, can reduce or minimize measurement error and bias. Texts on how to conduct good surveys exist (see, *e.g.*, Mitchell and Carson), and the techniques are not described here.

Reliability and Measurement Error

Reliability measures the variability among responses; valuations with relatively low variation among responses are considered more reliable estimates of value. Reliability measures whether the responses are consistent with each other, and thus is comparable to precision in statistics; it does *not* measure whether the responses accurately estimate the true value of the good. (The latter measure is called validity, and is discussed below.) Reliability contains a deterministic component (the normal variation in values among individuals) and random error due to imperfections in the survey instrument and/or sampling variance.

Sampling procedures, which are controlled by the researcher, influence reliability. Two approaches can improve the reliability of responses: (1) larger sample sizes, and (2) robust statistical techniques to handle "outliers" (responses that are considered too extreme, relative to the presumed distribution). Robust statistical techniques adjust responses that do not represent the true value (*e.g.*, a response of "99" to "number of dependents in the household") and would significantly influence the total valuation. Average willingness-to-pay, for example, can be significantly influenced by very high responses. Using robust techniques, such as the median value or a trimmed value (eliminating a percentage of responses from both ends of the distribution), generally improves the statistical reliability of contingent value surveys.[35]

Validity and Bias

Validity measures how accurately the contingent valuation of the good estimates the good's true value to society. This is comparable to accuracy in statistics, and bias is the term for the difference between the estimated value and the true value. Four categories of validity can be used to assess whether the responses are biased: (1) content validity; (2) criterion validity; (3) construct validity; and (4) theoretical validity. Contents are typically deemed valid, if the survey questionnaire is unambiguous and accurate, and closely matches the theoretical concept to be

---

[34] Portney, p. 8.

[35] Mitchell and Carson, pp. 211-229.

measured. Since questionnaire surveys are necessarily subjective, content validity is always a concern with such surveys.

Criterion validity requires comparison with some other method that is closer to being theoretically accurate, such as an estimate based on a derived demand curve. For goods where use is the majority of the value, prices from simulated markets can be used for comparison; an example would be a simulated market for hunting permits. For goods where nonuse accounts for the majority of the value, hypothetical values can be compared to actual referenda results.

Construct validity compares different measures for consistency. One form is to compare two methods, such as contingent valuation and the travel cost method, to see if the results are reasonably consistent; this, in essence, assesses the correlation between two or more measures. Significant differences, such as were found for the *Exxon Valdez*, raise questions about the construct validity of the contingent valuation survey.[36]

Theoretical validity evaluates whether the results are consistent with theoretical expectations; this typically involves a regression of the willingness-to-pay with other, independent variables to check whether the direction, magnitude, and strength of the relationships among variables are consistent with what would be expected under economic theory. The lack of criteria and truly comparable methods makes some of these tests of the validity of contingent valuation difficult, but surveys can usually be evaluated for their content and theoretical validity.[37]

Certain aspects of contingent valuation surveys could influence responses and lead to biased results. Biases can arise in numerous ways, because individuals be-have differently in various settings. Respondents may interpret the questionnaire differently, may be motivated by different aspects of the scenario when making decisions, may respond based on inferences about the use of their answers, or may use different cost-minimizing procedures or rules-of-thumb to make decisions when they know little about the good. Mitchell and Carson describe four types of biases:[38] (1) incentives to misrepresent responses; (2) implied value cues; (3) scenario mis-specification; and (4) sampling design and benefit aggregation biases.

Incentives to Misrepresent Responses. Compliant and strategic behaviors may lead respondents to inaccurately represent their preferences, because there are no incentives to tell the truth when the constructed market is hypothetical. Com-pliance

---

[36] On the other hand, in the case of the *Exxon Valdez*, the contingent valuation survey tried to measure the total value of the losses caused by the oil spill, while the travel cost method tried to measure the lost recreational value. One might anticipate that the total value would greatly exceed the recreational value, because: (1) the beauty and uniqueness of the Alaskan coast are well known, but the distance to Alaska inhibits recreational use, thus making nonuse values substantial, relative to use values; and (2) the recreation increase in the second year might be attributable to "rubbernecking" that occurs with many disasters, and does not necessarily offset the nonuse value losses of the disaster.

[37] Mitchell and Carson, pp. 189-209.

[38] Mitchell and Carson, pp. 231-293.

bias occurs when respondents give answers that they feel the interviewer wants. Surveying by a neutral party can usually correct for such bias, but respondents may still feel the need to give a "right" or "normal" (*i.e.*, compliant) answer.

Strategic bias arises when respondents intentionally misrepresent their preferences, because they believe it will influence the amount of the good provided, the amount or system for collecting money to provide it, or in damage appraisals, the compensation. Table 3 shows the types of strategic behavior, depending on the likelihood of the good being provided and the perceived obligation to pay for the good.

Table 3. Strategic Behavior in Valuing Public Goods

	Obligation to pay perceived as the amount offered	Obligation to pay perceived as being uncertain	Obligation to pay perceived as being fixed
Provision of good perceived as contingent on revealed preference	True preference (reveals true value)	Variable (true value might be overstated or understated)	Overpledge (overstates true value)
Provision of good perceived as likely, regardless of revealed preference	Free ride (understates true value)	Free ride (understates true value)	Nonstrategic minimum effort (answers that minimize time/ effort)

Source: Mitchell and Carson, p. 144.

The table describes predictions of how individuals would act under different payment and provision characterizations. A contingent valuation survey is intended to identify true values. True values are most likely to be revealed when both the fees charged and the amount provided will be based on the response (*i.e.*, on the stated willingness-to-pay). On the other extreme is minimal effort, where respondents feel they will have to pay a fixed amount and the good will be provided regardless of what they say. The other categories contain different types of strategic behavior that will cause respondents to "bid" inaccurately. Free-riding (underbidding) is more likely to occur when respondents feel that the good will be provided irrespective of their response. Overpledging (overbidding) is more likely to occur when respondents believe that the good is more likely to be provided with higher bids, but the bidders expect to pay a fixed amount, regardless of the bids. Most contingent valuation studies fall into the variable category because the payment amount is typically uncertain and provision is usually believed to depend on stated amounts; in this situation, free-riding and overpledging are both possible outcomes.

Other individual behaviors may mitigate strategic behavior, including altruistic motives, personal honesty and integrity (interest in telling the truth), the belief that many people are being interviewed in the survey, consideration of one's budget constraint when offering a bid, and the possibility that the good may not be provided at all. Nonetheless, if respondents have beliefs about the likelihood of the good being

provided or about the obligation to pay, or if they infer such information from the survey, strategic behavior can bias the valuation.

Implied Value Cues. Another type of bias arises when respondents decide on a valuation based on some particular aspect of the survey. This characteristic of the survey appears to give them a clue as to the "right" answer even if it were not intended to do so. *Starting point bias*, which typically occurs when using a bidding game format, can result when respondents feel that the starting bid is intended to approximate the correct value. A related problem — "yea-saying" — can occur when respondents simply accept a bid, even if it doesn't match their true valuation. The payment card approach was developed to correct the starting point bias of bidding games.

Respondents also infer values from other aspects of the survey. For example, some respondents give high valuations, because they feel that a study would not be conducted unless the resource or project being valued was important. Some methods, such as payment cards, include a range of values and typically benchmarks to suggest how much is spent on other (presumably similar) commodities. *Range bias* can occur when respondents' valuations are higher or lower than the highest or lowest amounts listed, when the amounts listed influence the bids, or when respondents do not find their valuations listed on the card. *Relational bias* can occur when respond-ents focus on benchmarks (particularly benchmarks related or similar to the good in question) to help determine their valuation. Finally, if several items are being valued, respondents may infer an indication of their values from their order in the list; typically, items listed first are perceived as being "more valuable" than items listed later. Thus, *position bias* could lead to invalid results. Altering the order in different interviews can overcome this bias, but it substantially increases the number of inter-views needed.

Scenario Misspecification. A third type of bias, scenario misspecification, can arise when the scenario is either not specified properly according to theoretical or policy information (*theoretical bias*) or it is interpreted incorrectly by the respondents (*methodological bias*). Theoretical misspecification occurs when part of the survey is incorrectly specified, based on theoretical knowledge or policy information; this bias can usually be minimized with sufficient research beforehand to check the survey's consistency with theoretical and policy guidelines.

Methodological biases can occur in numerous ways, including[39]
- when respondents value the symbolic nature of a good, rather than the amount (resulting in the same willingness-to-pay for different levels of the good);
- when respondents include items beyond the level of the good in question, such as items outside of the specified location and benefits often associated with (but not part of) the good in question;
- when respondents use a different measurement scale (*e.g.*, general qualitative terms rather than exact numerical changes) than the researcher intended;

---

[39] Mitchell and Carson, pp. 231-259.

- when respondents are skeptical that the good will be provided, that adequate funding exists, or that the project will achieve the desired goals after completion;
- when respondents value the good differently based on how it is funded or who provides it;
- if respondents fail to adequately reconcile purchases with their income constraints;
- when respondents give an amount that they think the project will cost, feeling that if they bid their higher valuation, a portion will be wasted;
- when respondents use other materials in the survey (such as general questions in the opening part of an interview) to help come to a decision; or
- if respondents don't treat unrelated valuations as independent (similar to position bias, discussed above).

The need to devise a realistic scenario leads to three criteria to assess the survey instrument: familiarity, understandability, and plausibility. Particularly important survey elements include: the description of the good, the quantity produced, the market, the payment vehicle, and the elicitation method. The scenario may be familiar to respondents, if they have previous information; if not, it must be easy to understand and must convey new information effectively. The scenario must convey the expected change and consequences accurately; the studies of Grand Canyon air quality and Glen Canyon Dam re-operation were criticized for presenting excessive quality improvements. The scenario must also seem plausible or responses may not be meaningful. The researcher faces a realism-bias tradeoff, because more informa-tion makes the scenario more realistic but may cause strategically biased responses. When respondents feel that the survey is unrealistic, they typically give "don't-know" responses, guess randomly, or respond to other cues. This is particularly a problem for assessing damages after a disaster (such as the *Exxon Valdez* oil spill), since the scenario for a contingent valuation cannot be an event that has already occurred.

Sampling Design and Inference Biases. The other principal type of bias arises when sampling design or benefit aggregation is not performed properly. The sample used for the contingent valuation survey must be designed so that the appropriate population is sampled and the sample fully represents that population. Determining the appropriate population (of individuals or households) can be difficult when the people who pay differ from those who benefit from the proposed change. This is further complicated by affected private property owners and individuals with exist-ence values for the good who live at substantial distances from the affected site. A sufficiently large population is needed to capture all of these values. Furthermore, the portion of the population that is sampled must accurately represent the values of the entire population, or the results could be biased. Inadequate or unrepresentative samples were criticisms of all three of the applications discussed earlier.

Nonresponse (to the survey in general or to specific questions) can also lead to biased results. Nonresponse to questions can include: don't know; refusal to answer; protest zeros (obviously erroneous answers, usually of zero, to register objection to the survey or the issue); and responses that are not internally consistent. Follow-up questions are usually necessary to distinguish between protest zeros (where respond-ents do not agree with the survey procedure) and actual zeros (where respondents would pay nothing for the good). Internally inconsistent responses (*e.g.*, responses

that are improbable or infeasible, given the identified income) can also bias results. Other outliers can be eliminated by statistical techniques or judgment, but arbitrary or biased decisions could affect the validity of the survey. Stratified sampling procedures can moderate nonresponse biases among distinguishable groups (where the individuals value the good differently) within the population, but cannot account for nonresponse biases among people with similar characteristics who have different nonresponse rates and value the good differently.

Inference biases may occur when the results from one particular contingent valuation study are used to estimate the value of a different good.[40] *Temporal selection bias* may occur when data from one study are used for a different time period, because public preferences for the good may change. However, evidence from two sources — public opinion polls and other contingent valuation studies — indicate that valuation results are fairly stable over time; valuations are also likely to be more stable for a good with which respondents are familiar than for one with which they have little experience.[41]

Sequence aggregation bias may occur when data from independent studies are aggregated over additional locations or goods. For example, if several areas are to be cleaned up, the valuations of each area measured in a particular sequence may differ from the valuations of each area measured independently or in a different sequence, because of income and substitution effects. Money "spent" on the first area in the survey typically reduces the amount identified as being "spent" on other areas (income effect) and the first area may act as a substitute to some of the features of additional areas to be valued (substitution effect). Items reached at a later point in a sequence thus are likely to be valued less than if they were valued independently or earlier in the sequence. Bias arises when values from these kinds of surveys are aggregated without considering the sequence of valuations.

Empirical Criticisms

Several studies have attempted to empirically assess the reliability and validity of contingent valuation surveys. One should be aware, however, that such studies typically use existing contingent valuation surveys or conduct new ones, and thus are subject to all of the errors and biases of the contingent valuation surveys being evaluated. Therefore, their conclusions may be no more reliable or valid than the results of the surveys they critique.

Diamond, *et al.*[42] Four researchers used their own contingent valuation studies to determine whether economic preferences are actually being measured. They focused on a criticism called the embedding effect — that willingness-to-pay is the

---

[40] Mitchell and Carson, pp. 261-287.

[41] Mitchell and Carson, pp. 261-287.

[42] Peter A. Diamond, Jerry A. Hausman, Gregory K. Leonard, and Mike A. Denning, "Does Contingent Valuation Measure Preferences? Experimental Evidence," in *Contingent Valuation: A Critical Assessment*, ed. J.A. Hausman (Amsterdam: North-Holland, 1993), pp. 41-62.

same whether one item is valued or several items are valued. This is similar to symbolic and sequence aggregation biases, discussed above (under Scenario Mis-specification and Sampling Design and Inference Biases, respectively). The example presented by Diamond *et al.* is that similar valuations resulted from different num-bers of wilderness areas protected. Proponents of contingent valuation argue that income and substitution effects explain the discrepancy in values. Diamond *et al.* counter that this effect is insufficient to explain the large variation in values observed in contingent valuation studies. Because the portion of income lost in valuing a sequence of goods is typically small, relative to average income, they conclude that income effects are insignificant.

Diamond *et al.* conducted several tests to determine whether substitution effects could explain differences in valuations. Other researchers have noted that, in a se-quence of valuations, the valuation of goods later in the sequence will be lower than the valuation obtained independently, if some items can be substituted for others. Diamond *et al.* designed a survey to test the hypothesis that respondents would be willing to pay higher income taxes to prevent the development of more wilderness areas. They posited that, as more areas are developed, fewer substitute wilderness areas exist for recreation, so the current area being considered should be valued more highly. Their results led them to reject the hypothesis, implying that the substitution effect is not large, at least in this case. Further tests were conducted to examine whether alternative means of measuring the same quantity yielded the same answer. For example, they compared the value assigned to two areas (with seven already developed) to the sum of the value of one of the areas (with seven already developed) plus the value of the other area (with eight already developed). Using parametric tests, which put less weight on outliers, they conclude that such different ways of measuring the same quantity fail to give similar results, and thus violate one of the validity standards described above. Diamond *et al.* argue that these results arise from a "warm glow" effect, where respondents feel a sense of improved well-being by contributing to a good cause, and that contingent valuation does not measure true economic preferences.

Desvousges *et al.*.[43] A study by six researchers was conducted to determine if contingent valuation surveys yield valid and reliable results. The authors used three hypotheses to test for validity and reliability. Data on willingness-to-pay to protect different numbers of migratory waterfowl by improving response services for oil spills were used to test these hypotheses. Based on a statistical analysis of the re-spondents, the researchers concluded that responses to different levels of protection were taken from the same population.

The first hypothesis was that higher levels of a good would elicit higher values. To test this hypothesis, the authors used an open-ended question to measure the value of protecting 2 thousand, 20 thousand, and 200 thousand migratory waterfowl from small and all oil spills. The results showed similar valuations across the changes in

---

[43] William H. Desvousges, F. Reed Johnson, Richard W. Dunford, Sara P. Hudson, and K. Nicole Wilson, "Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability," in *Contingent Valuation: A Critical Assessment*, ed. J.A. Hausman (Amsterdam: North-Holland, 1993), pp. 91-114.

quantities, leading the authors to reject the hypothesis and conclude that contingent valuation surveys were not valid.

The second hypothesis was that open-ended and dichotomous-choice questions would yield similar results when used to value the same quantity. To test this hypothesis, the two formats were used to measure the difference in the value associated with differing levels of response service for oil spills. The authors found that the dichotomous-choice format yielded a significantly larger number of high bids and generally yielded higher results than did the open-ended questions. Since the questions were measuring the same quantity but yielded different results, they rejected the hypothesis and again concluded that contingent valuation does not yield valid results.

The third hypothesis was that the results would not be affected by the procedures used to handle the data (such as functional forms or the bid structure), to assess the reliability of contingent valuation results. The first test compared total values calculated using linear and nonlinear functional forms for responses. The second compared total values from a survey using a high bid of $250 versus another survey using a high bid of $1000. The authors found that results varied significantly, leading them to reject the hypothesis and conclude that estimates from contingent valuation surveys are not reliable, as well as not valid.

Kahneman and Knetsch.[44] These two researchers concluded that responses to contingent valuation questions represent people's willingness-to-pay for moral satisfaction rather than for the good in question. They also concluded that people derive more benefits when they contribute more to a good cause, rather than when they consume more. The authors found that a ranking of projects based on moral satisfaction predicts the ranking by different willingness-to-pay amounts with a high degree of correlation. Willingness-to-pay, as an index of moral satisfaction, also helps to explain the embedding effect discussed by Diamond *et al.,* because addi-tional amounts of the good may add little to moral satisfaction. The second point made by the authors is that many individuals have a portion of their budget already devoted to purchasing moral satisfaction. They found that measured willingness-to-pay for additional moral satisfaction reduced discretionary spending, rather than reducing (substituting for) current purchases of moral satisfaction.

Conclusion

Contingent valuation is becoming more widely used in natural resource damage appraisal and in decisionmaking. It is and will likely remain controversial, however, because it is a complicated and imperfect device. Its application is an expensive and time-consuming research project, and a host of potential problems make the results of contingent valuation surveys suspect. However, the relevance or magnitude of the many types of errors and biases described in this report can only be assessed for each

---

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

survey; it is impossible to reach an unqualified conclusion as to the reliability and validity of such surveys generally.

When attempting to assess public preferences, nonuse values are real, and at times significant, possibly exceeding use values substantially. Proponents contend that excluding nonuse values in calculating damages and in decisionmaking would understate total values affected, and that contingent valuation is a theoretically valid way to estimate nonuse values. Opponents argue that the methodology is weak and the measures are not comparable to traditional measures of utilitarian values (because resource use generates economic and social benefits beyond those measured by price and volume), and thus can lead to arbitrary assessments of damage. Congress has recognized such values in directing federal land management agencies to "balance" values produced and protected. Congress has more explicitly acknowledged nonuse values in damage recovery programs, and may debate methods for measuring such values, including contingent valuation, particularly in any consideration of reauthorization of Superfund