Using Surveys to Value Public Goods: The Contingent Valuation Method

Robert Cameron Mitchell ■ Richard T. Carson

Resources for the Future
Washington, D.C.

Third Printing 1993

Printed in the United States of America

Published by Resources for the Future
1616 P Street, N.W., Washington, D.C. 20036
Books from Resources for the Future are distributed worldwide by The Johns Hopkins University Press

Library of Congress Cataloging-in-Publication Data

Mitchell, Robert Cameron.
Using surveys to value public goods: the contingent valuation method

Bibliography: p.
Includes index.
1. Public goods—Valuation. 2. Public goods—Cost effectiveness. I. Carson, Richard T.
II. Title.
HB846.5.M58 1989 363 87-28633
ISBN 0-915 707-32-2

The book was edited by Samuel Allen and designed by Peggy Friedlander.

♾ The paper in this book meets the guidelines for permanence and durability of the Committee on Production Guidelines for Book Longevity of the Council on Library Resources.

RESOURCES FOR THE FUTURE

RESOURCES FOR THE FUTURE (RFF) is an independent nonprofit organization that advances research and public education in the development, conservation, and use of natural resources and in the quality of the environment. Established in 1952 with the cooperation of the Ford Foundation, it is supported by an endowment and by grants from foundations, government agencies, and corporations. Grants are accepted on the condition that RFF is solely responsible for the conduct of its research and the dissemination of its work to the public. The organization does not perform proprietary research.

RFF research is primarily social scientific, especially economic. It is concerned with the relationship of people to the natural environmental resources of land, water, and air; with the products and services derived from these basic resources; and with the effects of production and consumption on environmental quality and on human health and well-being. Grouped into four units—the Energy and Natural Resources Division, the Quality of the Environment Division, the National Center for Food and Agricultural Policy, and the Center for Risk Management—staff members pursue a wide variety of interests, including forest economics, natural gas policy, multiple use of public lands, mineral economics, air and water pollution, energy and national security, hazardous wastes, the economics of outer space, and climate resources. Resident staff members conduct most of the organization's work; a few others carry out research elsewhere under grants from RFF.

Resources for the Future takes responsibility for the selection of subjects for study and for the appointment of fellows, as well as for their freedom of inquiry. The views of RFF staff members and the interpretation and conclusions of RFF publications should not be attributed to Resources for the Future, its directors, or its officers. As an organization, RFF does not take positions on laws, policies, or events, nor does it lobby.

1

Valuing Public Goods Using the Contingent Valuation Method

Our national commitment to a cleaner and safer environment has persisted in the face of oil embargos, stagflation, concerns about economic competitiveness, and competing budgetary claims. But as we progress toward the goal of a cleaner environment, each successive improvement becomes more costly to accomplish than its predecessor. Given finite public resources and restless taxpayers, this inevitably raises some difficult policy questions. How clean should we make the air? Should we attempt to make the lower Mississippi River as pure as the lakes in Wisconsin? Just how high a level of impurity should we tolerate in our drinking water? Is a further expansion of a state's park system justified in the face of the legitimate needs of industrial developers? Do we buy another B1 bomber? How much more do Medicare recipients value access to their traditional doctor than being enrolled in a health maintenance organization? Economists believe that questions like these can be addressed with empirical research in the form of benefit-cost analysis. By balancing the costs of public goods against their benefits, decision makers can arrive at more informed choices, or so the logic goes. In recent years the demand for such an accounting has found increasing favor among federal and state policy makers.

Unfortunately, few endeavors are more difficult than assigning a dollar value to something as elusive as increases in air visibility, or keeping the option of paddling a canoe in a wilderness preserve. Economists have long measured the value of goods that are routinely bought and sold in the marketplace. But ordinary markets do not exist for "public" goods such as national defense, the Apollo program to send man to the moon, and many environmental amenities.[1] Sometimes, as in the case of recreation sites,

[1] Pure public goods are characterized by the conditions of non-excludability of and non-rivalry congestion between individuals who wish to use the good (Cornes and Sandler,

this is because under public policy there is no charge for the good or service, or there is an arbitrarily determined charge (which does not reflect the full cost of providing the service or its true market value). In other cases, such as air and water quality improvements, which are public goods in the truest sense of the word, a charge would not be feasible because once the amenities are provided, people cannot be excluded from enjoying them.

For decades economists have grappled with the challenge of valuing public goods. The contingent valuation (CV) method is one of a number of ingenious ways they have developed to accomplish this demanding and important task. For reasons to be presented throughout this book, we argue that as things now stand, contingent valuation represents the most promising approach yet developed for determining the public's willingness to pay for public goods. Generally speaking, it appears as accurate as other available methods, it requires the researcher to make fewer assumptions, and it is capable of measuring types of benefits that other methods can measure only with difficulty, if at all. Our message is one of optimism tempered with realism. Like all sophisticated methodologies, contingent valuation presents challenges, and an important focus of the book is on the pitfalls of using the method. Contingent valuation's use of surveys to obtain consumer responses to hypothetical situations makes it vulnerable to various types of error, which we consider in detail so the researcher can take steps to avoid them and so the policy maker can evaluate and use CV findings with confidence.

The Contingent Valuation Method

The CV method uses survey questions to elicit people's preferences for public goods by finding out what they would be willing to pay for specified improvements in them. The method is thus aimed at eliciting their willingness to pay (WTP) in dollar amounts.[2] It circumvents the absence of markets for public goods by presenting consumers with hypothetical markets

1986). They may be seen as a special type of externality. In the real world, few public goods meet these strict conditions; we discuss the implications of deviations from this definition for contingent valuation in chapter 3. Whether something is a "good" or a "bad" depends on one's perspective. For example, an increase in environmental quality may be a "good" to consumers, and a "bad" to producers who happen to bear the immediate cost of pollution control.

[2] Respondents to CV surveys may also be asked what level of compensation they would be willing to accept (WTA) for a loss, but in what follows we refer to the WTP format unless otherwise noted. For reasons elaborated in later chapters, the WTA format should usually be avoided in CV studies because it does not elicit valid data under many circumstances.

in which they have the opportunity to buy the good in question. The hypothetical market may be modeled after either a private goods market or a political market. Because the elicited WTP values are contingent upon the particular hypothetical market described to the respondent, this approach came to be called the contingent valuation method (Brookshire and Eubanks, 1978; Brookshire and Randall, 1978; Schulze and d'Arge, 1978).[3] Respondents are presented with material, often in the course of a personal interview conducted face to face, which consists of three parts:

1. *A detailed description of the good(s) being valued and the hypothetical circumstance under which it is made available to the respondent.* The researcher constructs a model market in considerable detail, which is communicated to the respondent in the form of a scenario that is read by the interviewer during the course of the interview. The market is designed to be as plausible as possible. It describes the good to be valued, the baseline level of provision, the structure under which the good is to be provided, the range of available substitutes, and the method of payment. In order to trace out a demand curve for the good, respondents are usually asked to value several levels of provision.
2. *Questions which elicit the respondents' willingness to pay for the good(s) being valued.* These questions are designed to facilitate the valuation process without themselves biasing the respondent's WTP amounts.
3. *Questions about respondents' characteristics (for example, age, income), their preferences relevant to the good(s) being valued, and their use of the good(s).* This information, some of which is usually elicited preceding and some following reading of the scenario, is used in regression equations to estimate a valuation function for the good. Successful estimations using variables which theory identifies as predictive of people's willingness to pay are partial evidence for reliability and validity.

If the study is well designed and carefully pretested, the respondents' answers to the valuation questions should represent valid WTP responses. The next step is to use these amounts to develop a benefit estimate. If the sample is meticulously selected by means of random sampling procedures, if the response rate is high enough, and if the appropriate adjustments are made to compensate for participants who fail to respond (nonrespondents) and for those who give "poor"-quality data, the results can be generalized with a known margin of error to the population from which

[3] At different times and in various places the contingent valuation method has been called the survey method, the interview method, the direct interview method, the direct questioning method, the hypothetical demand curve estimation method, the difference mapping method, and the preference elicitation method.

pure public goods.[4] Other established methods for measuring benefits, as we will see later, tend to be limited in scope to the class of quasi-private goods. In the next section, we define the concept of benefits and show in more detail how they might accrue, employing a pure public goods example—national freshwater quality—to illustrate the various categories of benefits. The concepts we advance apply to, or are adaptable to, most pure public and quasi-private goods.[5]

The Nature of Benefits

The benefits from the provision of public goods such as pollution control arise from the value individuals assign (Brown, 1984) to improvements in air visibility or water quality and the like. Losses caused by a deficiency of pollution control may be thought of as damages. The distinction between benefits and damages is based upon some reference level of the good, a concept closely related to that of property rights. For ease of exposition, the term "benefits" will be used here to refer to both benefits and the reduction of damages. While the other side of the benefit-cost equation—the "costs"—may also have many meanings, we will try to avoid confusion by defining costs solely in terms of the value of the resources used to produce a public good such as pollution control.[6]

Each type of good has a unique configuration of advantages to confer on an economic agent (individual or household); expressed in another way, there are unique reasons why an agent's level of satisfaction or utility might increase if the good is provided to him. According to the economic doctrine of consumer sovereignty (Penz, 1986), an agent's spending behavior in markets is a sufficient signal of his preferences for various goods, and the reasons why he holds these values are of no economic importance. The question of why someone might value changes in the provision of nonmarketed public goods, however, is of more than passing interest to the researcher who wishes to measure benefits. The researcher needs to identify all the agents who might potentially benefit from a given change in order to have that population represented in the analysis. The identification of agents

[4] Methods similar to contingent valuation are used by marketing researchers to trace out demand curves for private goods—particularly new goods for which market data are unavailable. See, for example, Pessemier (1960) or Jones (1975).

[5] We will often use the term public goods when we do not wish to make a distinction between pure public goods and quasi-private goods.

[6] What we mean by costs is clearly conveyed by the term "control cost function," which is defined by Just, Hueth, and Schmitz (1982:278) as "the cost incurred by a polluter (due to reduced output and other opportunity cost in addition to direct pollution abatement cost) to reduce pollution from the level which would occur in a free (unregulated) market in the absence of control policies." The term hints that the calculation of the cost side of a benefit-cost analysis is not as straightforward as it might first appear.

is fairly straightforward conceptually, but often somewhat more difficult to accomplish in practice.[7] It is also important to identify the full range of possible benefits so that a method (or methods) can then be chosen which offers, all other considerations such as cost being equal, the most promise of measuring those benefits.[8] As we will show later in this chapter, the ability of different techniques to estimate categories of possible benefits varies. Where certain categories of benefits cannot be measured, they should be explicitly identified so policymakers can understand the limitations of the benefit assessment. Finally, an awareness of why and how agents benefit from changes in public goods can be helpful in knowing where to look for traces of the demand in actual private good market transactions, or, in the case of CV studies, in knowing which possible benefit categories to remind people to consider when they determine their willingness to pay for the good.[9]

A Typology of Possible Benefits

A comprehensive assessment of the benefits of a change in the level of a public good should include all of the benefits which will legitimately accrue from a specified change in the provision of a given good. This concept is sometimes known as the "total value" approach (Randall and Stoll, 1982; Boyle and Bishop, 1985). Some kinds of benefits are easier to measure than others, and the failure of economists to measure nonuse environmental amenities has long been criticized by noneconomists. Much of the history of benefits measurement can be written in terms of how researchers have devised ways to measure a larger and larger fraction of the total benefits of providing a public good (Smith, 1986a).[10] The CV method figures importantly in this history because one of its advantages is that it allows the researcher to directly measure the various types of nonuse

[7] For a discussion of problems in defining the appropriate population for a benefits study, see chapter 12.

[8] The need for an exhaustive compendium of benefits is well recognized by market research practitioners who study the potential market for new products (see, for example, Myers and Tauber, 1977).

[9] Agents benefit from the perceived changes in the public good rather than the actual technical changes. Credibility problems occur when actual conditions are more favorable than perceived conditions. The most cost-effective course of action by public officials in such cases would be to raise utility by an educational effort rather than proceeding with additional technical changes (Pauly, Kunreuther, and Vaupal, 1984). See Desvousges, Smith, and McGivney (1983) for a discussion of the chain of effects and responses to water quality regulatory actions.

[10] Unmeasured benefits have often been referred to in the economics literature as "intangibles" (Wyckoff, 1971; Bishop and Cicchetti, 1975; Haveman and Weisbrod, 1975), a term which reflects the propensity of applied economic analysis to define reality in terms of observable market behavior. A more appropriate designation for the unmeasured benefits is Ciriacy-Wantrup's (1952) concept of "extra-market" benefits.

return to a rural setting for wages far lower than a cost-of-living differential would suggest (Deaton, Morgan, and Anschel, 1982). In the cases of both existence value and psychic income, people are influenced in part by preferences for attributes of situations or goods that are secondary to the attribute which is presumed to be the primary source of value, such as the salary paid to a worker or the recreational use of a water body. In neither instance do the benefits result from the process of consumption as it is usually described in economic models, where the commodities are exhausted or used up (Smith 1986a).

Influential early discussions of existence benefits were presented by Krutilla (1967) and Krutilla and Fisher (1975). As the CV method has gained acceptance and benefits estimates have appeared that reflect a significant existence component, a number of economists have sought to clarify the nature of these benefits,[15] and others have attempted to use CV surveys to obtain separate measurements of one or more of the various types of existence benefits (see chapter 12). Despite the vigorous debate and profusion of terminology that characterize this literature, we believe the issues are now sufficiently well understood to permit CV researchers to measure with confidence the total benefits respondents receive for goods which include nonuse components.

What kinds of benefits might people obtain from the provision of a public good, apart from their personal use of the good? In figure 3-1 we identify four types of benefits, which fall under two categories—vicarious consumption and stewardship. In the case of vicarious consumption values, utility is gained from knowing about the consumption of others. These "others" may be generalized, or they may be particular individuals known to the respondent. The motivation behind vicarious consumption values may stem either from a sense of obligation to provide the good, or from a sense of true shared and interdependent utility. In practice, it is often difficult to get respondents to distinguish between these motives,[16] but for the purposes of economic analysis (including the design of a CV survey) it is rarely necessary to do so.

Stewardship values involve a desire to see public resources used in a responsible manner and conserved for future generations (Pigou, 1952; Ciriacy-Wantrup, 1952). We distinguish two types of stewardship values.

[15] For recent discussions, see: Bishop, 1982; McConnell, 1983; Brookshire, Eubanks, and Randall, 1983; Randall and Stoll, 1983; Smith, 1983; Freeman, 1984a; Edwards, 1985; Madariaga and McConnell, 1985; Brookshire, Schulze, and Thayer, 1985; Brookshire, Eubanks, and Sorg, 1986; Smith, 1986a, 1986b; and Hanemann, forthcoming. Driver, Brown, and Burch (1986) discuss the range of motives which may lie behind existence values for wilderness preservation.

[16] Gift-giving, altruism, philanthrophy, and interdependent utility are discussed from various perspectives in Winter (1969), Krebbs (1970), Schall (1972), Becker (1976), Goldman (1978), Mitchell (1979a), Margolis (1982), Sugden (1982, 1986), and Edwards (1985).

Bequest values exist when someone enjoys knowing that the current provision of an amenity will make it available for others—family or future generations—to enjoy in the future. Also part of stewardship are inherent values, which stem from the respondent's satisfaction that an amenity itself—a wilderness area, for example—is preserved regardless of whether it will ever be used by anyone. When people, in the absence of any intention to see harp seals in Canada or any expectation of doing so, contribute money to prevent the harvesting of baby harp seals for their pelts, they are manifesting something close to this type of existence value for that good.[17]

The different types of existence benefits coexist with each other and with use values. For example, while stewardship values do not result from current human use, they may be stimulated by and occur simultaneously with use: someone's stewardship value for wilderness lakes is likely to be enhanced by the experience of fishing in them during wilderness hiking expeditions. Thus, while the several dimensions of existence value are analytically distinguishable, and all enter into a consumer's utility function, they are likely to be very difficult to disentangle and measure separately. The same may be said for certain types of use values, such as the aesthetic (visibility) and health benefits from improved air quality.

While our definition of the existence class of benefits includes inherent values, Brookshire, Eubanks, and Sorg (1986) have recently argued that these values should be excluded from benefit measurements on the grounds that this type of motivation is not consistent with the "efficiency ethic" they believe underlies benefit-cost analysis. They describe the efficiency ethic as a management ethic focused on human welfare. For them, vicarious consumption and bequest values pass the efficiency-relevant test because these values involve efficient use of the resource in the interest of humans, but they do not believe this is true of inherent values. They argue that even though a person may be willing to pay something "simply because he believes we ought to protect wetlands wildlife against human action which would threaten the existence of the wildlife," this is not an acceptable motivation for benefits because it does not contribute to human welfare (Brookshire, Eubanks, and Sorg, 1986:1515). Actions based on ethical considerations, on a desire to do what is right, are "counterpreferences" since they do not increase utility. Brookshire, Eubanks, and Sorg advise researchers to probe individuals' motives underlying statements of willingness to pay in order to avoid confounding efficiency-relevant values and ethical values.

[17] The line between existence values of this kind and use values is blurred when, for example, an interest group attempts to provide a market in which one can purchase photographs and lobbying in behalf of protection of harp seals.

and $90, respectively. He can rank the cards optimally so as to minimize the disutility of being forced to make such a choice; this is the behavior expected by the researcher.[58] Note, however, that forcing the respondent to make a choice that he would not voluntarily make lowers his initial level of utility, thus violating the assumption underlying Hicksian compensating surplus.[59] Thus the desired unique inverse Hicksian demand function cannot be identified, and at best what can be traced out is a family of indifference curves possessing particular curvature properties.

Hypothetical/Direct Methods. By directly measuring people's valuation of particular hypothetical changes in amenity quality and quantity, Hypothetical/Direct measurement methods shortcut the need to make the large number of assumptions required by the indirect linkage methods. Hypothetical/Direct methods, which include contingent valuation, posit what Smith and Krutilla (1982) term "institutional" links between amenity levels and individual behavior. The institutional assumption is that individual responses to hypothetical markets are completely comparable with individual responses to actual markets, an assumption whose implications we consider at length in the following chapters. Once this premise is granted, however, methods become available which are unique in their simplicity, theoretical justification, and ability to value the entire range of benefit categories.

Since we have already described the contingent valuation method, the discussion of Hypothetical/Direct methods here will be limited to the spend more-same-less type of survey question and to allocation games with the possibility of a tax refund. The spend more-same-less survey approach is based on ordinary survey questions which ask respondents to say whether "we" (the United States) are spending too much, too little, or about the right amount for certain types of government programs (National Opinion Research Center, 1983).[60] As a result of interpreting the trichotomous responses to mean that "spending too much equals a preference for less of the good and lower payments," and so on, researchers have used incompletely developed scenarios as the basis for estimating demand curves for one or more public goods. An early and somewhat simplistic example of the economic analysis of such spend more-same-less questions is Akin, Fields, and Neenan (1973). More recent studies of this

[58] If the respondent regards the offered choices as so far from his preferences that he is unwilling to make the effort to optimally rank the disutility of the various alternatives, his ordering will be meaningless.

[59] Contingent ranking based on observed data (for example, a single choice and a set of possible alternatives) does not suffer from this problem because it is based on voluntary choice.

[60] The spend more-same-less survey question method is sometimes called the "micro-based estimates of demand function" approach.

type have employed increasingly sophisticated discrete-choice statistical techniques (Gibson, 1980; Bergstrom, Rubinfeld, and Shapiro, 1982; Gramlich and Rubinfeld, 1982; Ferris, 1983; and Langkford, 1985).

The obvious risk of the typical spend more-same-less survey question is that the analysis of preferences is based on superficial and uninformed responses. Compared with that in a contingent valuation scenario, the description of the good is sparse. For example, the commonly used question in the National Opinion Research Center's General Social Survey (1983) asks respondents to express their views about "solving the problems of the big cities" and "improving and protecting the environment." There is little effort to create a market or specify a payment obligation. Respondents are rarely given information about what is currently being spent on these programs and how it is being spent. Particularly problematic is the assumption that the respondent knows what his payment obligation would be for each response. Langkford's work (1985) casts doubt on the validity of this assumption.

We include allocation games that offer respondents the possibility of a tax refund among the Hypothetical/Direct methods because this type of allocation game allows the respondent to reject payment in favor of a tax refund for any and all public goods being considered, instead of merely requesting the respondent to allocate a fixed budget among different public goods. An advantage of this type of allocation game is that it forces the respondent to value a large number of goods simultaneously, rather than valuing each in isolation (as is typical of a CV survey). The disadvantages include the superficial description of the different categories of public goods and the fact that if no tax refund is desired, the WTP estimates obtained may not be maximum WTP estimates. The two best known examples of allocation games with tax refund are Strauss and Hughes (1976) and Hockley and Harbour (1983).

The Advantages of the Contingent Valuation Method

The advantages that the hypothetical methods in general, and Hypothetical/Direct methods in particular, offer to the benefits researcher are highlighted by the comparisons shown in table 3-3. Table 3-4 compares the four classes of benefit measurement methods according to five criteria—ability to measure option price, to value goods not previously available, to estimate all existence-class benefits, and to directly estimate the relevant ordinary and Hicksian inverse demand curves.

Of particular importance is the flexibility of the hypothetical methods. As Sen (1977:339–340) has observed, "once we give up the assumption that observing choices is the only source of data on welfare, a whole new world opens up, liberating us from the informational shackles of the

Table 3-4. Key Properties of the Benefit Measurement Methods

Desirable properties	Method			
	Observed/Direct[a]	Observed/Indirect	Hypothetical/Indirect	Hypothetical/Direct
Able to obtain option price estimates in the presence of uncertainty	No	No	Yes	Yes
Able to value goods not previously available	Yes	No	Yes	Yes
Able to estimate all existence class benefits	Yes	No	Yes	Yes
Relevant ordinary (or inverse demand) curve is directly estimable	Yes	No	No	Yes
Relevant Hicksian compensated demand (or inverse demand) curve is directly estimable	No	No	No	Yes

[a] In some cases, only referenda have the desired property.

traditional approach." Within the important constraint that the scenario must have plausibility for the respondent, the CV researcher can easily specify a variety of states of the good to be valued and the conditions of its provision. Moreover, these need not be limited to current institutional arrangements or levels of provision. Brookshire and Crocker, in reference to the inflexibility of methods based on observed behavior, point out that

> the only really sound way of obtaining an estimate of whether the net benefits of a particular property rights restructuring are positive, if one insists upon employing observed everyday behavior, would be to perform the restructuring and observe the results. In some circles, this is simply known as trial and error. Trial and error can be an extremely costly way to perform research because the errors are real rather than hypothetical. (Brookshire and Crocker, 1981:246)

The hypothetical character of contingent valuation, which allows it to obtain ex ante judgments, also permits it (and the other hypothetical-based methods) to obtain WTP amounts that include the existence values, whereas methods that rely on observed behavior can obtain existence values only with great difficulty, if at all. Consider the different values that two people might put on maintaining the current level of air visibility at Mesa Verde National Park. Person A values only that level of visibility which is current during his visit to the park. He is willing to pay some amount to enjoy this attribute during his visit. Person B has no current-use value for the air visibility in the park, but believes that national parks are an important part of America's heritage and represent a public obligation for which he is willing to pay. If a travel cost study of air visibility were made in Mesa Verde National Park, its methodology would assume that travel behavior reveals the respondent's price for the site's services, and that air visibility is jointly supplied along with the site's other attributes. Thus, such a travel cost study might be able to indirectly estimate person A's use value for the park's air visibility, but could not measure the stewardship values held by person B. A contingent valuation study, in contrast, is not hindered by this restriction. The respondents in a CV survey put a price on an amenity after assessing the total change in welfare that it represents to them. If the respondents are reminded of the relevant benefit categories that should be kept in mind,[61] their WTP amounts

[61] This is necessary to ensure uniformity in the respondents' conceptualization of the valuation situation. If they are not so reminded, some respondents may fail to consider nonuse values, either because they did not think of such values in the context, or because they mistakenly assumed that the interviewer wanted them to consider only the use dimension. Desvousges, Smith, and McGivney (1983) were the first to use a "value card" as a visual aid in interviews in order to remind respondents of the potential benefit categories. Under the headings "use," "might use," and "just because it's there," their card listed reasons why the respondent might value clean water in the Monongahela River.

should reflect the full range of their preferences. Thus in a CV study the WTP amounts of persons A and B would include both use and existence benefits (to the extent that they exist).

The Hypothetical/Direct methods (including contingent valuation) also are able to directly measure specific points on an individual's compensated demand curve. This avoids the problem, which plagues the other methods, of potential bias resulting from wrong assumptions about the form of the individual utility functions. For example, instead of imposing separability conditions on individual behavior, based on the researcher's assumptions about consumer tradeoffs, CV studies can allow the respondent to make his or her own tradeoffs in terms of money (Brookshire and Crocker, 1981:246).

Summary and Conclusions

In this chapter we have considered the nature of benefits and the wide variety of methods economists have to measure them. Public goods were divided into two types, quasi-private goods (public goods with individually held property rights) and pure public goods (public goods with collectively held property rights). This distinction has implications for the type of market model appropriate for use in designing a contingent valuation survey. The benefits of both kinds of public goods result from the values individuals place on these goods. These values are subjective and multidimensional. A correct benefits assessment will include all the benefits that legitimately accrue to a given improvement. An inventory of freshwater quality benefits was used to illustrate the principal types of benefits, which include use and existence benefits and their several subdivisions. This inventory did not include option value, on the grounds that option value is not a meaningful benefit category for ex ante welfare measures.

Our discussion of benefit measurement methods focused primarily on the relationships among the various types of behavior-based methods. We classified sixteen different benefit measurement methods into four types according to whether the method relied on preferences revealed in observed or hypothetical markets and on a direct or indirect linkage between the method and the willingness-to-pay value. The Hypothetical/Direct methods, which include contingent valuation, were shown to be the only class of methods simultaneously capable of obtaining option price estimates in the presence of uncertainty, valuing goods not previously available or marketed, estimating all existence class-benefits, and obtaining in a direct fashion the relevant Hicksian demand curves. It was the prospect of realizing these advantages that motivated the CV pioneers to explore the uncongenial (to the well-trained economist) realm of hypothetical markets.

4

Variations in Contingent Valuation Scenario Designs

Having described the relationship between the contingent valuation method and other methods of measuring the benefits of nonmarketed goods, we now turn to three key issues in the design of contingent valuation scenarios that are the subject of current debate among CV researchers. Each involves design choices the researcher must make that may have a large effect on the quality of a CV study's findings and their appropriateness for use by policy makers. The first is whether a CV study should be based on a private goods market or a political market. The second issue concerns which of the several techniques for eliciting the WTP amount from the respondent is most appropriate for the study, and how the chosen technique should be implemented. These two issues are discussed in this chapter. The third set of choices involves how much and what kind of information about the amenity and the hypothetical market the researcher should include in the material presented to the respondent during the course of the interview. These decisions involve tradeoffs on the researcher's part between the need to inform the respondent about relevant features of the hypothetical market and the need to avoid information overload, and between the desire to measure benefits in a manner that offers policymakers the utmost flexibility in using the findings and the difficulty respondents have with scenarios that are too abstract (in the sense that they lack concrete details about the amenity and the conditions under which it would be provided). Discussion of the information issue begins in this chapter and continues throughout the book.

Private Goods Markets and Political Markets

Contingent valuation studies simulate a market for a nonmarketed good. Until recently researchers took it for granted that a private goods market

13

Conclusion

The Promise of Contingent Valuation

Contingent valuation shows promise as a powerful and versatile tool for measuring the economic benefits of the provision of nonmarketed goods. It is potentially capable of directly measuring a broad range of economic benefits for a wide range of goods, including those not yet supplied, in a manner consistent with economic theory. Other available methods, in contrast, are capable of measuring only some of those benefits and are limited to valuing existing goods and existing quantity and quality levels, and researchers employing them must make a number of unverifiable assumptions in the course of deriving benefit estimates from the available data.

But can CV surveys actually measure values that are sufficiently reliable and valid for use in benefit estimation? Our conclusion is basically affirmative. Certainly the prophecy that respondents will usually act strategically and will deliberately bias their values is not supported by the evidence reviewed in chapters 6 and 7. Our assessment in this regard is similar to that reached by Arrow (1986:183), who states: "Neither the empirical evidence nor the theoretical arguments convinced me that strategic bias is liable to be significant (in CV studies)." Nor is the hypothetical character of CV surveys necessarily an impediment to their usefulness. Unlike ordinary surveys, which often measure unconstrained attitudes toward vaguely defined goods, CV surveys elicit specific behavioral expectations—for example, "I would be willing to pay an additional $X a year in higher taxes out of my current income in exchange for the specified improvement in local air quality." On the basis of our reading of the literature on attitudes and behavior, we believe that the properties which have been found to maximize behavioral predictions are largely compatible with the fundamental structure of most contingent valuation scenarios. Moreover, com-

parisons of the outcomes of identical simulated and hypothetical markets, in the controlled experiments reviewed in chapter 9, have shown that hypothetical markets for quasi-private goods are able to predict market outcomes when real payments are involved. These data lie behind the emerging consensus of skeptics (Kahneman, 1986, and Freeman, 1986, for example) and practitioners that CV studies are able to measure meaningful values for "familiar" goods such as local recreational amenities.

What about the use of contingent valuation to measure the benefits of less familiar goods, such as air quality improvements or risk reductions of various kinds? This is more difficult terrain because here we cannot test the accuracy of CV surveys against a criterion. Air visibility cannot be sold in simulated markets in the way that Bishop and Heberlein sold deer-hunting permits for Wisconsin's Sandhill preserve. Those who harbor the most qualms about the contingent valuation method's ability to obtain meaningful values for pure public goods hold the view that CV surveys should replicate a consumer market in familiar goods. According to this notion, the method can only work when respondents either have well-defined preference orderings for the goods being valued at the point that they are asked to take part in a study (see Freeman, 1986; Department of the Interior, 1986:27721), or when the valuation procedure provides an extended learning process by which firm preference orderings can be acquired; otherwise, in these views, valid WTP amounts cannot be obtained. Those who hold the latter view doubt that even a 45-minute personal interview could provide such a learning experience.

While sympathetic to the concerns raised by these skeptics, we believe their view that meaningful valuation requires previously "well-exercised" preferences is based on an inappropriate market model. In our view, the appropriate model for CV surveys of pure public goods—goods that citizens are least likely to have direct experience in valuing—is the referendum, by which citizens make binding decisions about the provision of public goods. From this perspective, instead of falling short of the relevant market model, well-conducted CV surveys offer significant improvements over actual referenda as a means of measuring consumer preferences.

First, CV surveys can elicit a more informed decision than referenda. Studies of voter decisions show that people typically face an informational vacuum when they vote on noncontroversial propositions, which leads them to rely on endorsements by political leaders or to make snap judgments (Magleby, 1984). In sharp contrast, relatively detailed and focused information is presented to respondents in the course of a CV interview. Second, CV surveys are based on a more representative set of responses than most referenda. In actual referenda those who vote are often not particularly representative of the entire voting population owing to low turnouts and the tendency of some categories of voters, such as those with low levels of education, to be underrepresented among those who do vote.

The use of probability sampling, callbacks, and imputation procedures makes it possible for CV surveys to reach a more representative audience.[1]

Provided that respondents can be motivated to carefully follow the contingent market described in the scenario, and find it sufficiently plausible, CV surveys offer the possibility of obtaining meaningful information about consumer preferences for nonmarketed amenities. During the course of the interview, respondents make a decision about how much they are willing to pay for the amenity based on the material presented in the scenario, any prior information they might have, and their preferences regarding what they would like the government to do with their tax dollars.

Relevance and Quality

Although the contingent valuation method is a promising technique, the fact remains that the methodological challenge in conducting a CV study is considerable because it is often difficult to convey to respondents what a policymaker wants them to take into account in a way that is both theoretically and technically correct *and also* understandable and plausible. This problem has often been underestimated by CV practitioners and ignored by benefit analysts, who tend to treat CV studies as if they were all of equal quality. But how can a potential user of a CV study know when to place credence in its findings? The answer is, only by assessing the study's relevance to the policy change the user wishes to value as well as the study's quality.

Relevance

As the number of contingent valuation studies increases, it is likely that studies whose scenarios were designed for one situation will be used to infer something about other situations. Evaluating a study's relevance or transferability for a particular purpose requires a careful mapping of the changes the study valued against the changes implied by the policy. One issue here is the correspondence between the provision of the amenity described in the scenario and the amenity changes implied by the policy changes. Another is the context in which the amenity was valued. In a CV study, the improvements that respondents value are specific to the scenario presented to the respondents. It is to be expected that the WTP amounts will be sensitive to the method of provision, the payment vehicle, and the other features of the hypothetical market used to elicit the values. The

[1] The real issue when using the political market model is whether the researcher desires to predict the relatively uninformed, unrepresentative vote which would occur if a referendum about the amenity valued in the CV survey was actually held, or whether the researcher prefers to predict what would happen if a relatively informed and representative vote occurred.