September 1991

Critique of Contingent Valuation and Travel Cost Methods for Valuing Natural Resources and Ecosystems

W. David Eberle

Boise State University

F. Gregory Hayden

University of Nebraska-Lincoln, ghayden1@unl.edu

Follow this and additional works at: https://digitalcommons.unl.edu/econfacpub

Part of the Economics Commons


This Article is brought to you for free and open access by the Economics Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Economics Department Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Critique of Contingent Valuation and Travel Cost Methods for Valuing Natural Resources and Ecosystems

W. David Eberle
and
F. Gregory Hayden

In July 1989 the U.S. District Court of Appeals in Washington, D.C. disallowed the method that had been established by the U.S. Department of Interior (DOI) for determining what a corporation would pay in the case of injury to an ecosystem from hazardous waste spills [Ohio v. Interior 1989]. The DOI method was based on neoclassical methodology and appraisal techniques. The Court ruled that "restoration is the proper remedy for injury to property where measurement of damages by some other method will fail to compensate fully for the injury," and that "natural resources have value that is not readily measured by traditional means" [Ohio v. Interior 1989, pp. 456–57]. One traditional means that the courts found would not compensate fully for injury is the utilization of market prices. As the court stated, "it is unreasonable to view market price as the exclusive factor, or even the predominant

The authors are Professors of Economics, Boise State University and Professor of Economics, University of Nebraska, respectively. This article was presented at the annual meeting of the Association for Institutional Thought, Portland, Oregon, April 1990. The authors wish to thank Theresa D. Hayden for her excellent research and administrative assistance on the U.S. Environmental Protection Agency research project of which the work found in this paper was a part. We also wish to recognize James A. Swaney for his thorough critique and extensive suggestions.

one. From the bald eagle to the blue whale and snail darter, natural resources have values that are not fully captured by the market system” [Ohio v. Interior 1989]. Thus “DO1 erred by establishing ‘a strong presumption in favor of market price and appraisal methodologies’ 51 Fed. Reg. 27, 720 (1986)” [Ohio v. Interior 1989]. While ruling against the dependence on market means for measuring injury from hazardous spills, the Court did not explicitly rule on the contingent valuation method (CVM) and the travel cost method (TCM) as appraisal methodologies for the valuation of natural resources and ecosystems. These methodologies, both of which are based on neoclassical ideology, attempt to place a market valuation on the natural environment that is not included in market exchange.

The purpose of this article is to critique CVM and TCM in order to assist courts and legislative bodies in future deliberations. It is divided into three sections. The first is to critique the CV and TC methods in the context of the neoclassical paradigm. This section also demonstrates that CV and TC methods reach to the same conceptual base as that used to legitimize market prices. The second section applies psychometric standards to evaluate the CV method because the CV approach employs surveys to determine the market price of nonmarket ecological components. The final task, contained in the third section, is to apply the principles of general systems analysis (GSA) to the CVM and TCM. Using GSA is consistent with the understanding that ecosystems are systems that conform to system principles. The GSA principles contained herein are consistent with institutional methodology.

CVM and TCM Critiqued With Neoclassical Principles

Several market type methodologies have been designed by neoclassicalists to value non-market goods. These methodologies have a common foundation. The fundamental assumption is that the value of all goods can be expressed in money equivalent terms and that value is based on a good’s utility to humans. If humans do not determine a use or exchange value for a good, then its existence is inconsequential. This assumes that all goods are created to serve man. This distinction becomes critical in valuing the natural environment. Once the overarching assumption is made that the basis for valuation is for humans, a number of assumptions follow for neoclassicalists. Many of the assumptions will be discussed below. Three basic ones are: that utility functions exist; that a utility function exists that can value non-market goods in rank order; and that these value rankings of non-market goods can be empirically identified by the TC and CV methods.
Contingent valuation is defined as “any approach to valuation of a commodity which relies upon individual responses to contingent circumstances posited in an artificially structured market” [Seller, Stoll, and Chavas 1985, p. 158]. Typically, this valuation methodology is used in situations where exchange value cannot be established through a market process. Thus, there is no price associated with these goods. The CVM utilizes a direct questionnaire approach to solicit individuals—responses that purport to reflect each individual’s valuation of a non-market good. The questionnaire attempts to simulate a hypothetical market for the good in question where the respondent indicates either willingness to pay (WTP) or willingness to accept (WTA) compensation for the non-market good [Bishop and Heverlein 1979, p. 926].

There are three primary issues surrounding this technique that create potential problems. The first is the operationalization of a utility function for estimating value. The second issue is the hypothetical nature of soliciting value rankings. The third issue is the question of validating the resulting data. The validation issue is one of determining if the questionnaire measures what it purports to measure. Richard Bishop and Thomas Heverlein reviewed the preliminary empirical results of their goose permit study and suggested, “when summed together these potential problems are sufficient to justify considerable skepticism about the accuracy of the resulting values estimates” [Bishop and Heberlein 1979, p. 926]. These sentiments are also reflected by Christine Seller, John R. Stoll, and Jean-Paul Chavas in their recreational boating study. They concluded that the CVM had problems with producing nonnegative demand curves and negative consumer surplus [Seller, Stoll, and Chavas 1985, pp. 172 and 75]. In other words, the contingent valuation responses to their open format questionnaire implied the respondents were willing to pay less for ramp fees than they actually spend during the recreational season.

The TCM attempts to estimate demand functions for non-market goods based on the notion of indirect costs. The method has been extensively used in the valuation of recreational sites [Seller, Stoll, and Chavas, 1985, p. 157]. The demand functions are estimated by using travel costs as a surrogate for value. In defining indirect costs, several significant assumptions are made. First, it must be determined what are substitutable sites or activities. Second, a decision has to be made on an appropriate value of time to travel to the site. Third, decisions have to be made on how to allocate the value of a site between its ambience and its various other activities. There are also a number of data requirement problems related to this approach. The editors of the book
Valuing Environmental Goods conclude that the problems in specification and data collection with the TCM "result in the dispelling of what was once regarded as the TCM's greatest potential strength: appealing to the notion that visitor values must equal or exceed travel costs" [Cummings, Brookshire, and Schulze 1986, p. 95]. It appears that the outcomes of the TCM are a function of the assumptions made.

Measurement models are designed to reflect a theoretical construct. Before measurement techniques such as CVM and TCM can be understood, the theoretical construct they purport to measure must be understood. In other words, what does a regression with these variables actually measure? What are the demand functions that economists claim drive the market system? Do they exist for non-market goods as well? These questions will be addressed below in discussions regarding consumer preference.

The hedonic price method (HPM) is not being covered in this article. However, since it is referred to later in discussions regarding the testing of CVM, it will be briefly defined at this time. HPM was introduced to split a good into its various attributes for the purpose of assigning values to particular attributes. Operationally in HPM, the commodity's market price is generally regressed against attributes in order to assign values to attributes. For example, a house has an overall function of shelter, but each house also has numerous attributes such as size, number of rooms, location, style, lot size, neighborhood, and so forth. "Estimation problems abound in efforts to implement the HPM—to name but two: persistent collinearity between 'important' variables and extraordinarily low explanatory power in regression equations" [Cummings, Brookshire, and Schulze 1986, p. 96].

Consumer Preference Theory

The concept that individuals rank commodities derives from consumer preference theory. There are two specifications of the demand functions in valuing non-market goods. However, in discussions about the validity of estimating a price for non-market goods, there is rarely any mention of the appropriateness of using either Hicksian or Marshallian demand functions as the theoretical foundation for what the TCM or CVM purport to measure. Rather, the discussion of CVM in the literature surrounds the psychometric issues of demand artifacts, internal validity, and self-generating constructs [Heberlein 1979; Seller, Stoll, and Chavas 1985; Thayer 1981]. The TCM struggles with specifying the model, given the data limitations, so that it remains consistent with theory [Desvousges, Smith, and McGivney 1983]. Thus, a com-
parison of CVM to HPM cannot provide evidence of the validity or accuracy vis-à-vis "true" value of CVM as a means of valuing public goods [Cummings, Brookshire, and Schulze 1986, p. 96].

The microfoundation of market demand is the theoretical construct of consumer preference. This construct is often referred to as consumer demand functions or "utility" functions. The utility construct develops axioms that represent how consumers value rank commodities in a market system. Confronted with a set of market prices and corresponding goods, the consumer can establish ex ante a ranking of preference for these goods based on the relative satisfaction the goods will yield to him. This ranking is known as the indifference map. Understanding the nature of how an individual establishes this ranking requires a review of the axioms underlying the decision. The axioms direct that consumer behavior is determined by logical necessity and not by the study of human behavior.

Neoclassical economics is highly axiomatic in that it describes behavior as conforming to classical mathematical logic. Without this conformity, neoclassical economic modeling falls apart. Thus, a series of axioms are assumed about how consumers rank their preferences. The basic assumptions, which are relaxed in more advanced analysis, illustrate the point that the assumed human behavior in the theory of the consumer is a highly artificial one. The axioms are as follows:

1. **Completeness**: For all \( x, y \) in \( X \) either \( x \geq y \) or \( y \geq x \) or both. This provides a complete ordering of preferences so that cycling of preferences does not occur. This axiom implies that two bundles can be compared.

2. **Reflexivity**: For all \( x \) in \( X \), \( x \sim x \). Trivial.

3. **Transitivity**: For all \( x, y, \) and \( z \) in \( X \), if \( x \sim y \), \( y \sim z \), then \( x \sim z \). This axiom is required if preference is to be maximized.

4. **Continuity**: For all \( y \) in \( X \) \{ \( x: x \geq y \} \) and \{ \( x: x \leq y \} \) are closed sets. It follows that \{ \( x: x \geq y \} \) and \{ \( x: x \leq y \} \) are open sets. This axiom rules out discontinuities.

5. **Strong Monotonicity**: If \( x \geq y \) and \( x \neq y \) then \( x > Y \). This axiom simply states that more is preferred to less.

6. **Strict Convexity**: Given \( x \neq y \) and \( z \) in \( X \), if \( x \geq z \) and \( y \geq z \), then \( tx + (1 - t)y \geq z \) for all \( 0 < t < 1 \). This axiom is the generalization of diminishing marginal return [Varian 1984, pp. 112-13].

These axioms purport to define consumer preference. From this set of axioms the familiar definition of utility functions emerges, with the characteristics of convex indifference curves and a single preference maximization, given a budget constraint. From the utility function the demand function is specified. If the notion of consumer preference as
defined in the neoclassical model is to be accepted, all the axioms must also be accepted. These axioms are necessary and sufficient for demand functions to exist. If one of these axioms does not hold, then the notion of consumer preference cannot be determined. For example, the continuity axiom is necessary to rule out certain discontinuous behavior. In other words, in the ranking of bundles of goods, if bundles are merged into new sets, the strict ordering of the original set must be preserved. If offering goods to the consumer in different combinations causes the individual to reorder his preference, as often occurs through marketing, the axiomatic logic fails and no longer can optimal solutions be derived. Consumers must behave in a "rational" manner. "Rational" behavior is defined by the axioms. It is possible to construct reasonable scenarios where each axiom may be violated in practice by consumers making decisions to purchase. As Hal R. Varian summarizes, "A utility function is often a very convenient way to describe preferences, but it should not be given a psychological interpretation" [Varian 1984, p. 112]. "It goes without saying that the axiomatic-deductive method has been in disrepute in recent decades, in all disciplines but mathematics and formal logic—and even here the axioms are often supposed to be a mere convention rather than necessary truth" [Rothbard 1979, p. 20].

Proponents of this logical approach to model building argue that the model may be valid even if the assumptions are not. If the model approximates behavior, it is a legitimate tool of empirical analysis. [Friedman 1953, p. 15]. However, consumer preference theory and the resulting demand curves also have significant deficiencies in explaining or predicting behavior.

Alfred Marshall, and later John Hicks, made the first real efforts to empirically measure demand curves. Marshall approached the problem by specifying a demand function \( x(p,y) \) where demand is a function of price and income. This is a cardinal demand function where, in real terms, the consumer is able not only to value rank, but also to perceive the magnitude of value ordering. This implies that the demand function should be operational. In other words, consumer utility is measurable in money. It should then be possible to aggregate utility.

However, there are two partial derivatives to this function: the price differential and the income differential. For the expenditure minimization problem to be empirically identified, the function must be separable, thus implying that price and income are independent of each other. Therefore, the cross partials are zero. The separability requirement also implies that it is separable only up to a monotonic transformation. This
Contingent Valuation and Natural Resources

severely restricts the form the demand function may take. Human behavior is being constrained by mathematical requirements. The limitations and specification problems of this approach are such that "cardinalism in the 'additive' sense is therefore found quite untenable, while in the 'measurable-up-to linear-transformation' sense it is found completely inconsequential, except in an extreme situation, which is demonstrably unreal" [Majumdar 1975, p. 135]. It should be noted that the TCM is usually specified as a Marshallian demand function [Seller, Stoll, and Chavas 1985, p. 162].

A Hicksian demand function is an ordinal ranking demand function. Hicks operated directly from the utility function $x(p, U)$. Instead of minimizing expenditures as Marshall attempted, Hicks sought to maximize utility. Hicks devised an indexing method to resolve the income substitution problems encountered by Marshall. Hicks assumed that he was able to create an index of all goods except one, the numeraire. In essence, this puts goods in what is called real terms. Or, price has no real impact on the consumer choice. Thus, the sum of the cross partials is equal to zero. In other words, Hicks resolved the substitution and income effect problem of optimization. He achieved this by monetizing the goods. He assumed that money is neutral (has no effect) on real goods: therefore, there can be no income effect from price changes. However, there are significant deficiencies in trying to operationalize the Hicksian demand function. The first issue involved in determining the validity of this construct is the assumption that money has a neutral effect. The commodity chosen as a numeraire will affect the outcome of the ordering [Arrow 1981, p. 142]. Also, the comparison between two goods using a numeraire can only be made when "the market clears." When this condition of equilibrium is not met, money matters [Arrow 1981, p. 140]. Thus, changes in relative prices and changes in the numeraire will change the outcome of a Hicksian demand function preference ranking.

When an economist specifies a demand function in an econometric model, the type and form of the demand function is based on the above axioms rife with their conceptual difficulties. What is it these models have measured? The results, the economic parameters, are interpreted as the functional magnitude of the demand function. Thus, in a linear model, the weighted variables' contribution to the slope of the demand function account for a certain percent of the variation. The balance is attributable to random error. What one cannot say is that this estimation represents the demand function. The construct cannot account for the functional dichotomy of substitution and income effects or any of
the nonmonotonic discontinuous complex behavior of people. In this approach human behavior is assumed to be monotonic and the econometric model will insure that the results conform to this requirement. The models force the data to fit a construct. What is important to gain from this discussion of demand functions is that they are incapable of modeling a complex system. Most economists agree that demand functions, if they do exist, have never been observed and probably never will [Varian 1984, p. 142].

The neoclassical debate on how to empirically measure demand functions continues. However, the debate does not question the basic axioms of consumer preference. The consumer preference axioms are integral to the neoclassical model’s ability to construct optimization solutions. These derive the familiar conclusions that competition is a pareto optimal solution in the allocation of resources. However, these axioms do not yield usable operational rules for understanding real world phenomena. The problems encountered by functional form, restrictive assumptions, and naive beliefs on how people value rank goods leave consumer preference theory in the world of abstract constructs.

Other approaches for rank ordering preferences have attempted to overcome the specification deficiencies of demand functions. One approach that has received attention is called revealed preference. This set of axioms is the basis of CVM. In the event that a demand function cannot be revealed, it is still possible to generate comparative statics, both sign and ordinal ranking of goods. The general axiom of revealed preference is, if \( x_1 \) is revealed preferred to \( x_2 \), then \( x_2 \) cannot be strictly revealed preferred to \( x_1 \). \( x_1 R x_2 \) implies not \( x_2 R x_1 \). This is the observable consequence of utility maximization in that if data satisfies this axiom, the axiom is sufficient to meet the consumer preference axioms. Again, what is so important to note is the nature of classical deductive reasoning. If the data meets the revealed preference axiom, it can be deduced that because this condition is sufficient to meet the consumer preference axioms, the consumer preference axioms must be true. Unfortunately, theoretical work in the area of game theory has shown this axiom is invalid where strategies of a game theory type can be employed [Majumdar 1975, p. 139]. Game theory is used where consumers perceive trade-offs in their decisionmaking. The conclusion resulting from these game theory exercises is that the presence of choice need not reveal preference. The distortion of preference in observed choice suggests that preference may be sometimes revealed and sometimes distorted [Majumdar 1975, p. 143]. In other words, interpretation of results cannot be made.
This uncertainty in solution implies that revealed preference fails one of the critical foundations of scientific analysis, the rejection of plausible alternative hypotheses. What other explanations also satisfy the revealed preference axiom? Have they all been scientifically rejected? If not, then the conclusion cannot be drawn that meeting the revealed preference axiom implies the existence of consumer preference as described by the consumer preference axioms. In contrast, the conclusion drawn by neoclassicalists is that if there exists a utility function that could have generated that behavior, then the demand function exists [Varian 1984, p. 143].

Neoclassical economists continue to search for a resolution to the problem of empirically identifying demand functions. Another significant problem yet to overcome is aggregation of consumer preferences. It is not possible to derive a social welfare value by simply taking the preference of one individual and multiplying it by the number of persons in the society as explained by Arrow’s impossibility theorem. The impossibility theorem suggests that a society cannot have all individual preferences met and achieve optimal social welfare simultaneously. This is a problem that all public policymakers constantly confront in creating legislation. There are always winners and losers when new rules are imposed on society. The theoretical solution to resolving the impossibility theorem is that one of the axioms must be violated. The debate often surrounds dropping the pareto condition or the libertarian condition [Sen 1976]. If the transitivity axiom is dropped, it may be possible to have optimization both of individuals and of society [MacKay 1980, p. 89]. But the transitivity axiom is a necessary condition of consumer preference theory.

Consumer Preference and Non-market Goods

The notion of using Hicksian and Marshallian demand functions for non-market goods adds further complications to identifying consumer preference. The TCM and CVM are two examples of attempts to measure the magnitude of value for non-market goods [Bishop and Heberlein 1979]. These approaches assume that people can and do make comparisons among all goods. If this is true, then it should be possible to construct a demand function for these goods. It is assumed that people will value all normal goods in the same manner, or that the consumer preference axioms hold for all normal goods. Nonmarket goods do not have a price. Therefore, we do not know how people would respond to the pricing of these goods.

The very nature of non-market goods suggests there is some attribute
that cannot be captured by price. Thus, there will always be controversy as to whether TCM and CVM understate or overstate the relative real value. For example, when valuing the distance a person travels to visit a recreational site, does the greater the distance driven imply this non-market good is of greater value or does the person value the act of driving? What is the opportunity cost of travel time? These are questions that cannot be answered by aggregating the cost of gasoline and auto depreciation. Using a questionnaire format to answer these questions only raises new methodological questions concerning the validity of the results. Additionally, since demand functions have only been theoretically defined for those goods that have competitive markets, the notion that an indirect price can be constructed for a non-market good will always leave the TCM and CVM methods open to question.

CVM and TCM Critiqued by the Application of Psychometrics

The CV method uses a questionnaire format to solicit a Hicksian preference value from individuals. The issue that needs to be addressed is whether this is an appropriate technique for determining non-market good valuation. This section will review the use of psychometrics and its techniques as a methodology for revealing individual preferences. It is important to understand the restrictions and limitations of this measurement tool.

In psychometric theory there are very specific rules in questionnaire development and evaluation. These rules help insure that the instrument (questionnaire) measures what it purports to measure. One of the primary aspects of standardization requires that different people using the same instrument or an alternative instrument measuring the same trait should obtain similar results [Nunnally 1978, p. 3]. In the effort to standardize instruments so that repeatability is possible, a set of criteria is established that, if met, would provide reasonable confidence in the accuracy of the results. A complete guarantee can never be achieved because the consumer value for a non-market good is an unobservable trait. The two rules relevant to the CVM are that a questionnaire should have both reliability and validity. Because of the great potential for systematic error in the questionnaire approach, the instrument must be carefully tested to generate the maximum confidence possible that it is measuring what it is intended to measure, and that the instrument measures the trait accurately.

Reliability relates to the internal structure of the instrument. In a traditional instrument there are a series of questions that purport to meas-
ure the same trait. Thus, by checking the covariance of the responses to each question (Cronbach’s coefficient alpha), one can identify whether or not the items covary. If significant covariance exists, the researcher can conclude that the items are measuring the same trait. It is assumed that any one question imperfectly measures the unobservable trait. Thus, a series of questions that imperfectly measure the trait are used so that an overall image of the trait emerges. In the absence of using the reliability coefficient, an alternative check of reliability is the test-retest method. In the latter case, the questionnaire instrument is given to the same sample twice. After the initial survey, the questionnaire is given again several weeks later to see if the results are similar. There are serious defects in this approach that bias the reliability either upward or downward. Respondents may remember their responses from the first interview, or they may have changed their minds in response to an event occurring in the interval between interviews [Nunnally 1978, Chap. 7].

The reliability of the CVM instrument is an open question. The nature of the bidding game precludes the ability to compute a coefficient alpha because the format is different from the traditional approach to measuring an unobservable. The test-retest method could be used replete with the problems inherent in the approach. Even with the inherent problems, it may provide some insight into whether the instrument is measuring consistently.

In most CVM studies, no significant attention is paid to reliability and, therefore, it remains unclear if the CVM values generated are repeatable. Instead, the researchers compare their results with alternative measures for the same population and the same non-market good. This verification technique does not resolve the issues of instrument reliability or accuracy. Cummings, Brookshire, and Schulze argue that “all of the comparison studies undertaken to date have failed to carefully assess the accuracy either of the CVM used or the accuracy of the HPM (or TCM) used for comparison” [Cummings, Brookshire, and Schulze 1986, p. 72].

The second consideration is the issue of validity. Several types of validity measures are used to verify whether an instrument is measuring what it purports to measure. The type of validity relevant here is construct validity. When construct validity exists, it improves the confidence that a correspondence between the abstract construct and the instrument exists. This is a core issue for the CVM. Construct validity is a necessary condition for theory testing. There are two aspects of construct validity that must be considered. The first is termed internal
(trait) validity. This is essentially a term to identify if the construct is something different from other constructs. Essentially, this check tries to identify if systematic variance exists within the instrument and if this variance results in high correlations with other measures of the construct, and low correlations with measures to which the construct should not be related [Peter 1981, p. 207].

There are several approaches to establishing trait validity. One method that has received attention is the multimethod multitrait matrix. This technique attempts to identify what is termed convergent and discriminant validity. Convergent validity refers to two instruments with maximally different methods of measuring the same construct that are compared to see if their results converge. This is analogous to comparing TCM with CVM. The results should be similar if they are measuring the same construct. Discriminant validity refers to using the same method for different constructs to determine if there is low correlation between the two constructs [Campbell and Fiske 1959].

This approach has been tried by researchers attempting to validate their results by comparing different approaches for the same construct. The TCM should yield similar results to the CVM. What the researchers in the area of valuation of non-market goods have failed to explore is whether or not using the same methodology for different constructs will result in similar values. Convergent and discriminant validity is necessary but not sufficient to validate the construct [Campbell and Fiske 1959, p. 106]. By failing to pursue all available means to verify the existence of a construct, “we all stand to drown in a mass of meaningless and potentially misleading junk” [Jacoby 1978, p. 87].

The second form of construct validity that must be considered is nomological (lawlike) validity. This form of validity is based on the explicit investigation of the constructs and measures in terms of a formal hypothesis derived from theory. Nomological validation is investigating both the theoretical relationship between different constructs and the empirical relationship between measures of those different constructs [Peter 1981, p. 135]. In other words, do the results of CVM confirm or contradict neoclassical theory? An affirmative answer is a necessary condition for the acceptance of the instrument. Given all the problems associated with trying to measure an unobservable, except through hypothesized rules of correspondence, researchers conclude that a single study cannot validate the construct. In addition, it is not possible to conclude that the instrument is measuring what it purports to measure. This is essentially a form of external validation. “Even tentative acceptance of construct validity requires some amount of aggre-
Contingent Valuation and Natural Resources

The first step in the evaluation of results including both logical and deductive reasoning and a series of reliability and validity studies ... In fact ... the most important implication of construct validation is the increased emphasis on the role of theory in validation" [Peter 1981, p. 135]. Results that contradict theory should be carefully considered before measure is accepted. For example, the systematic error in CVM between the willingness to pay (WTP) and willingness to accept (WTA) measures is a case of unresolved construct validity. R.D. Willig [1976] argued that price changes and A. Randall and J.R. Stoll [1983] argued that quantity changes will have small income effects. Therefore, the substitution effect is observable. This conclusion is consistent with the theoretical constraints of the Hicksian demand functions so that WTP and WTA should be close for a given individual. The empirical results, however, demonstrate that WTA is consistently larger, on the order of three to five times larger, than measures of WTP [Cummings, Brookshire, and Schulze 1986, p. 35].

Another potential problem in using questionnaires arises in the areas of demand artifacts and self-generated validity. Demand artifacts include all aspects of the experiment that cause the subject to perceive, interpret, and act upon what he believes is expected or desired of him by the interviewer. These artifacts can take the form of the "faithful" subject providing answers that the subject believes the researcher wants, or the subject may take a negative or apprehensive role [Sawyer 1975, p. 20]. The CVM questionnaires have attempted to reduce this effect by fully explaining the process to gain the trust of the subject and to discard what are perceived as protest bids. A protest bid is the case in which the response of an individual is not consistent with that of his peers. The demand artifact also has been noted to be biased by what has been termed "starting point bias," that is, the subject perceives the relevant range of acceptable bids by the level where the bidding process begins, and thus, the results are biased [Cummings, Brookshire, and Schulze 1986, pp. 29–34].

A more serious demand artifact problem in the hypothetical market technique is trying to measure a trait that does not already exist in long term memory. In the CVM, respondents are asked to value non-market goods in terms of dollars. It is likely that many of the subjects have never considered this exercise before. What may occur is that the questionnaire becomes directive in how the subject will respond. How the market is presented to the subject may determine his beliefs about the market. This is especially a problem when the subject is required to value natural resources or an ecosystem the respondent does not under-
stand and has usually never observed. Having no foundation in his memory, the presented market becomes the basis for his decision [Feldman and Lynch 1988, p. 424]. What is likely to occur after the interview is that the respondent may well further reflect on the exercise and change his mind. Again, CVM is aware of the information bias [Cummings, Brookshire, and Schulze 1986, p. 33], but proposes that a more complete education will reduce the problem. In the case of self-generated validity, the respondent must already have considered the valuation problem and drawn conclusions to avoid this type of bias.

While psychometrics and econometrics are similar in attempting to test refutable hypotheses, there is one significant difference between them. In psychometrics, when the hypothesis is rejected, the researchers must consider whether the model is misspecified, the data is biased, or whether the construct is invalid. In econometrics, the construct (neoclassical foundation) is not questioned. The model may be misspecified, the data may be biased, or the sample may be inadequate. But hypothesis testing does not refute theory.

Empirical Studies: Travel Cost Method

The studies that have used the TCM have met with limited success. The problems with model specification and data limitation have biased the results so that the authors themselves have seriously questioned the validity of the conclusions. These studies are premised on the assumption that demand functions exist for non-market goods. Most of the studies use a Marshallian demand function. The results derived from these demand functions contradict the a priori predictions of the theory. The most common explanations given for the failure of these studies are two significant problems that confront the researchers. The first problem is the additivity of individual demand functions across sites and across activities to arrive at an aggregate value. The second related problem is the severe data limitations that restrict the specification of the demand functions to keep them consistent with theory.

The William H. Desvousges and Kerry V. Smith study on the value of water quality improvement for the United States Environmental Protection Agency (EPA) addressed the additivity and data limitations of this approach [Desvousges and Smith 1984]. The results of this study were similar to an earlier study on water quality for the EPA conducted by Desvousges, Smith, and Matthew P. McGivney [Desvousges, Smith, and McGivney 1983]. The conclusion of the 1983 study was that the explanatory power of the models used was not high [Desvousges,
Smith, and McGivney al. 1983, pp. 1–10]. In the 1984 study, the authors attempted to improve the model specification of the demand function. As a validity check on their results, both Hicksian and Marshallian demand functions were specified in order to compare the results with theory.

Desvousges and Smith were forced to assume that the individual utility functions were equal for all individuals for each activity to resolve the additivity problem [Desvousges and Smith 1984, pp. 2–21]. The model requires monotonically separable demand functions but, due to data limitations, the authors were forced to assume an average mean user as a representative individual. The results were disappointing to the authors, who recognized the ad hoc nature of their assumptions [Desvousges and Smith 1984, pp. 2–28]. The crudeness of these assumptions seriously affects the quality of the results.

In many instances the estimated coefficients did not agree a priori with the expected signs [Desvousges and Smith 1984, pp. 6–38]. Theory predicts that the Marshallian consumer surplus should be greater than the Hicksian surplus. The results contradicted theory [Desvousges and Smith 1984, pp. 7–21]. In an attempt to explain the results, the authors suggested that one significant contributor to the poor results was the use of the proxy variable, the mean user. The mean user included a mix of activities undertaken at a site, which is clearly inconsistent with theory [Desvousges and Smith 1984, pp. 6–38]. Additionally, the authors recognized that the variables specified, and the valuation assumptions made, make a substantial difference in the final benefit estimates [Desvousges and Smith 1984, pp. 8–24]. For example, each individual's valuation of the opportunity cost of travel time to a site is unknown. For some people travel time is perceived as a form of recreation, while for others it is time lost from work. These different valuations cannot be summed to a meaningful aggregate. Thus, the researcher is forced to assume a value for travel time that is difficult to verify as accurate.

The Meta Systems report prepared for the EPA reached similar conclusion. [Meta Systems 1987]. They argued that the values generated should not be taken as important or precise in themselves. In other words, they are only approximations. They believed their calculated values were a result of their assumptions. The researchers believed that their assumptions were conservative and therefore underestimated the true value of consumer surplus [Meta Systems 1987, pp. 1–25]. When their TCM results were compared with the CVM results, they failed to converge. The authors emphasized that these results underscore the limitations and shortcomings of these methodologies [Meta Systems
Contingent Valuation and Natural Resources

little better than the TCM studies. While similar problems exist for the CVM with respect to using utility functions, these studies introduce an additional source of potential error through the use of hypothetical markets and questionnaire format for data collection.

Whereas most of the attention in the studies using the TCM focused on the model specification and data limitation problems, the CVM studies focused on the problems of validity. These models, in constrast to the TCM, primarily used the Hicksian demand function that assumes utility remains constant when calculating monetary value for the non-market goods. This approach assumes away the additivity and separability issues but still contains the state-dependent problems.

In constructing a hypothetical market, the researchers are confronted with several tasks. The first is determining how to reveal the value of the goods. This involves educating the respondents as to the nature of the hypothetical market. To get the respondents to valuate these non-market goods, two approaches have been tried: WTP for using the good and WTA for not using the good. Theoretically, these two valuations should be approximately the same, assuming no income effect. The results of the research that used both approaches found that there was a significant divergence between the two measures. The WTP tended to undervalue the asset, while the WTA tended to overvalue the asset. The over and under valuation is based on the valuations in relation to each other, since there is no market valuation of the goods [Bishop, Heberlien, and Kealy, 1983, p. 620]. In this sense, we do not know if the results of the two measures are under, over, or otherwise skewed. These results contradict the a priori theory that suggests that the two measures should be approximately the same given the predicted small income effect. This lack of convergence is particularly worrisome. The foundation of the utility function hypothesis is based on R.D. Willig's assertion that the two measures should be synonymous [Willig 1976]. To avoid this a priori contradiction, many studies “have not even bothered to estimate WTA” [Cummings, Brookshire, and Schulze 1986, p. 137].

A second primary issue with the CVM approach is how to solicit the WTP valuation from respondents. Most authors have chosen a bidding process in which an initial price for the non-market good is offered. The respondent then indicates if the bid is low or high. The bid is then adjusted accordingly until the respondent feels his value on the non-market good is reflected. This method has received much analysis and criticism for what is called starting point bias. If the opening bid is too low, then it appears that this will bias downward the respondent's final choice of value. The converse is also true. If it starts too high, it will bias upward the final bid [Boyle, Bishop, and Welsh 1985, p. 193]. The
study done by Seller, Stoll, and Chavas used both an open-ended format in which the respondent provided the valuation, and a closed-ended format in which the respondent answered yes or no to the stated value [Seller, Stoll, and Chavas 1983]. Revealing in this exercise were the substantially different results that occurred. The authors concluded that the open-ended format may be unreliable because of the negative consumer surplus and the low results this methodology produced [Seller, Stoll, and Chavas 1985, p. 175]. However, an alternative explanation could account for the low results of the open-ended questionnaire and the relatively better performance of the closed-ended questionnaire. First, it may be that individuals who have not valued the good in question in a market-oriented setting may be unable to place a value on the non-market good. Second, the close-ended questionnaire may have a starting point bias, or may be creating self-generated valuations on the basis of how the information is presented.

The self-generated construct is one of the most significant issues confronting the CVM. "The specific valuation problem may be so remote from the respondents' market valuation experiences as to leave him unable to respond reliably" [Brockstael, McConnell, and Strand 1988, p. 25]. Thus, several researchers advocate either educating the respondent sufficiently that he is able to respond intelligently [Thayer 1981, p. 38], or interviewing only those who have proximity and therefore, knowledge of the non-market good [Beasley, Workman, and Williams 1986]. Either approach introduces a bias that diminishes the reliability of the results. Additionally, the second case ignores the existence value of the non-market good—that is, those who will never use the good, but gain value from knowing that it is there. Cummings et al. conclude that the CVM may yield accurate values where respondents have made actual choices for the good in a market framework. Their review of the CVM studies does not show that people are capable of making the valuations the CVM is asking of them [Cummings, Brookshire, and Schulze 1986, p. 102].

Like the TCM results, the CVM results have generally been disappointing. The studies caution the reader that the results do not have high enough internal reliability or generalizability to justify conclusions [Smith and Desvousges 1987; Desvousges et al. 1987b; Brockstael, McConnell, and Strand 1988; Bishop and Heberlien 1979; and Seller, Stoll, and Chavas 1985]. The errors most commonly identified by the authors are hypothetical bias, strategic bias, information bias, and interviewer bias [Cronin 1982, p. ix]. There is considerable debate on the significance of each of these biases and discussion on how to minimize each bias. These discussions usually surround why the studies failed to pro-
vide significant results or why they contradicted a priori theory predic-
tions. For example, in the Kerry V. Smith and William H. Desvousges
[1987] study on the value of risk changes, the results rejected the hy-
pothesis of a declining marginal valuation of risk with reductions of
the risk level [Smith and Desvousges 1987, p. 109]. The authors ob-
serve that this seems to contradict rational behavior. The Cronin study
in 1982 identified respondents engaged in strategic voting [Cronin
1982, pp. 6-10]. This implies that true preferences are not being re-
vealed by the respondents. Bishop and Heberlien, in reviewing the bias
of their goose permit study, suggest that "when summed together these
potential problems are sufficient to justify considerable skepticism
about the accuracy of the resulting values estimates" [Bishop and
Heberlien 1979, p. 926].

The CVM has been critically reviewed in a recent book by Ronald
G. Cummings, David S. Brookshire, and William D. Schulze [1986].
The editors focus on what they consider to be the core issues sur-
rounding CVM. The first issue is the sources of bias in the CVM. The
second issue is the accuracy of the method. The bias issue is at the foun-
dation of their debate in that, if substantial bias exists, it casts serious
doubt on the validity of the method. The authors emphasize the need
to focus on the range of divergence between the WTP and WTA. They
also urge that a standard for what is an acceptable range consistent with
a priori theory (starting point bias) should be established. Two causes
have been identified for this divergence between the WTP and WTA.
The divergence may occur as a result of how the respondent is to make
his offered payment (vehicle bias) and the lack of incentive for accurate
valuation because the respondents are spending hypothetical dollars
(hypothetical bias) [Cummings, Brookshire, and Schulze 1986, chap. 3].

After the studies were reviewed for accuracy, eight were chosen. When comparing the results of CVM with another indirect market
measure (TCM), the hypothesis that the measures were the same was
rejected [Cummings, Brookshire, and Schulze 1986, p. 105]. Rejection
of the hypotheses resulted from lack of familiarity with the non-market
good, lack of experience in valuing the non-market good, uncertainty
in the choice structure, and lack of resolution between WTP and WTA
criteria.

Many of the authors still cling to their belief that CVM can be used
in public decisionmaking. They argue that the biases may be identified
or mitigated, and that the difference between actual versus hypothetical
payment is weak [Cummings, Brookshire, and Schulze 1986, pp. 146
and 242]. However the authors agree that the CVM is most likely to
succeed where the operationalization of the hypothetical market
occurs—in other words, where participants are familiar with the non-market good and have experience in valuating the good and also understand the hypothetical market valuation method. When these conditions hold, it is also the case that there are other methods available in addition to CVM.

**Appropriateness Conclusion**

The results of the travel cost method and contingent valuation method have generally been unverifiable and burdened with significant operationalization problems. Cummings, Brookshire, and Schulze conclude their review by saying that the “CVM may not be as hopeless as we and others earlier believed. ‘Promise’ is not ‘performance’, however, and our assessment refers only to the potential promise of the CVM as a viable method for estimating values for public goods. The realization of that promise implies real challenges for theoretical and empirical research” [Cummings, Brookshire, and Schulze 1986, p. 234]. Their conclusion about TCM as a viable method is even more pessimistic. “The environmental (and other public good) ‘commodities’ for which the TCM or RPM might be used for valuation purposes are very limited, however” [Cummings, Brookshire, and Schulze 1986, p. 6].

Cummings et al. fail to raise the fundamental issue of the appropriateness of the theoretical framework used for these models. Regardless of the efforts to make the models more sophisticated, or to create new data sources, doubt remains as to whether these methods will produce results that have meaning.

**CVM and TCM Critiqued with GSA Principles**

The adverse criticism related above regarding CVM and TCM was from within the neoclassical paradigm, as articulated in the neoclassical literature for the neoclassical context. This criticism, however, does not apply to the systems context. In this section CVM and TCM will be compared to General Systems Analysis (GSA) principles. The CVM and TCM methodologies are inconsistent with GSA, and are not an attempt to define or evaluate a system. Next, GSA principles will be defined and used as standards by which to judge the adequacy of the CV and TC methodologies. Any methodology is adequate for some context. For example, systems of mathematical mind games can be judged to be adequate in the context of mind games. The task here is to judge methodological adequacy with respect to a real-world system context.
System Defined

“A system is a set of objects together with relationships between the objects and between their attributes” [Hall and Fagen 1968, p. 81]. Objects are the elements and components of the system. Attributes are the properties of the elements and components, and relationships are what tie the system together. The relationships to be considered “depend on the problem at hand, important or interesting relationships being included, trivial or uninteresting relationships excluded” [Hall and Fagen 1968, p. 82]. To use Kenyon B. De Greene’s definition, “in the most general sense, a system can be thought of as being a number or set of constituents or elements in active organized interaction as a bounded entity, such as to achieve a common whole or purpose which transcends that of the constituents in isolation” [De Greene 1973, p. 4].

There is no end to a system. Any relationship or delivery between components leads to additional deliveries, and to positive and negative feedback deliveries. One-dimensional systems (such as would be implicit in an assumption that human consumption is the end of the economic system) are not real world systems.

The main GSA criticism of neoclassical valuation is that one aspect, market demand, of one system element, the human agent, is being used for system evaluation in a manner that treats the human element as the only user of the ecosystem. A fundamental principle of the modern systems approach is to avoid the analysis of an entity in isolation. To overcome the dangers inherent in the human need to categorize the universe into separate entities, “every system must be analyzed within the context of its environment” [Mattessich 1978, p. 21]. However, the context of CVM and TCM is not a system within an environment; their context is utility analysis.

Utility is a more fundamental problem to these techniques than the isolation problem just mentioned, because utility does not exist in the real world. Therefore, neither does a utility function exist. This is readily made explicit by most economists. As Hal R. Varian has stated, “a utility function . . . should not be given a psychological interpretation” [Varian 1984, p. 112]. Lionel Robbins, a respected participant in the pure theory of utility analysis has stated that it “has had a perennial fascination for some of the best minds on the subject” [Robbins 1975, p. ix]. However, he clarifies that “the pure theory of value is not one of those branches of economic analysis which have any immediate bearing on practice” [Robbins 1975, p. ix]. The economist Tapas Majumdar, in his book on the Measurement of Utility (which is not about trying to measure utility in a real world sense), states that “on more
than one occasion in the preceding pages, we have made the observa-
tion that the nature of welfare perception in the physiological or psy-
chological sense is necessarily beyond the scope of our discipline" 
[Majumdar 1975, p. 32]. Herbert A. Simon explained in his Nobel Prize 
recipient lecture that on the basis of numerous studies, the idea that 
people behaved so as to maximize subjective expected utility (SEU) 
was false.

The refutation of the theory has to do with the substance of the decisions, 
and not just the process by which they are reached. It is not that people 
do not go through the calculations that would be required to reach the 
SEU decisions—neoclassical thought has never claimed that they did. 
What has been shown is that they do not even behave as if they had carried 
out those calculations, and that result is a direct refutation of the neoclas-
sical assumptions [Simon 1979, p. 507].

The social and psychological sciences rejected the idea of utility 
around the turn of the century. After hedonism, utility, and instinct 
theory fell into disrepute, an attempt was made in social psychology to 
substitute other reductionist ideas. These, however, fell into disrepute, 
and psychology came to the same conclusion as GSA—that individual 
beliefs, attitudes, and tastes are the result of an integrated system, not 
arising from individualistic natures, or hedonistic urges, or utility. 
[Harre 1983].

The reductionist approach, which attempted to leave social welfare 
to utility calculation, was also denied by the historical tide. The tide 
turned toward government policy to protect and enhance social and 
economic welfare. The claims of utility calculation and hedonism 
“when tested in the crucible of social policy, proved inadequate” [All-
port 1985].

It is worthwhile to know of the origins of the idea of a utility function 
in order to understand why economists as well as psychologists and 
social psychologists have rejected the idea. Philip Mirowski has ex-
plained its origin in a number of articles. It came from an energy for-
mula of mid-nineteenth century physics, not from experiments or 
observations of humans. The utility function was “asserted to represent 
a gravitational field, which by the 1860s was also identified as the field 
of potential energy. This is why Tjalling Koopmans can state that “a 
utility function of a consumer looks quite similar to a potential function 
in the theory of gravitation” [Koopmans 1957, p. 176]. “The metaphor 
of energy utility which was appropriated by neoclassical economics was 
derived from the physics of a specific historical moment, namely, the 
years of the mid-nineteenth century just prior to the elaboration of the 
second law of thermodynamics” [Mirowski 1987, pp. 84–85]. There
was no empirical base or even introspection that would indicate human rationality should be defined "as the maximization of an objective function over a conserved entity" [Mirowski 1987, p. 84]. Instead, our economic ancestors were quite honest that they were borrowing the physical metaphor to render consumption theory a mathematical science. "Jevons (1905b, p. 50), Walras (1960), Edgeworth (1881), and nearly every other early neoclassical economist admitted this fact" [Mirowski 1987, p. 83].

In any case, the concept of utility is a unidimensional concept that assumes that human utility maximization is the end or purpose of the system. Thus it ignores the elements, constituents, components, and their relationships, and therefore the concept is inconsistent with GSA.

Openness

All real-world systems are open systems, and all open systems are non-equilibrium systems. "Open systems are those with a continuous flow of energy, information or materials from environment to system and return" [De Greene 1973, p. 36]. There are misconceptions that arise both in theory and practice when social organizations are regarded as closed rather than open. "The major misconception is the failure to recognize fully that the organization is continually dependent upon inputs from the environment and that the inflow of materials and human energy is not a constant" [Katz and Kahn 1976, p. 101]. Systems and their environments are open to each other, as are subsystems within the systems. Living systems both adapt to their environment and modify their environment.

GSA divides the analysis between the system under consideration and its environment. The system description is referred to as the internal description, or the state of the system. However, all systems are influenced by an external description outside the boundaries of the system. An example is the work of EPA's Environmental Monitoring and Assessment Program (EMAP). EMAP has found that a wetland ecology receives inputs such as contaminants, sediment, and nutrients from agriculture ["Environmental Monitoring..." 1989]. Although inputs from (often called forcings) and outputs to (often called responses) the external environment are important to the system, no attempt is made to define the environment's structure. It only has a functional "black box" description to the system. The term environment as used in systems analysis may mean an ecosystem, for example, if the system under study is a socioeconomic system. If the system of interest is an ecosystem, then the socioeconomic system is the environment. This concept is displayed in Figure 1.
Figure 1. Conceptualization of Open Systems
In systems analysis, environment refers to the functional area outside the system. Because real world systems are constantly open to their environment, they cannot reach an equilibrium state. It is one of the goals of analysis to match up the two kinds of system descriptions. "The external description is a functional one; it tells us what the system does, but not in general how it does it. The internal description, on the other hand, is a structural one; it tells us how the system does what it does" [Rosen 1972, p. 53].

Four external functions of the natural environment for the social system have been defined by D.W. Pearce in his Environmental Economics and refined by James A. Swaney [1987]. The functions are:

1. Natural goods production, which includes wilderness, greenery, landscape, scenery, and so forth. It is often competitive with natural resource production, and is restricted in quality and quantity by the production of effluents from households and production centers.

2. Natural resources, the raw material and energy sources flowing from the environment, upon which the production of goods and services is dependent. Natural resources represent only part of one of the two flows from the environment to the economy. They flow to the private and public production centers.

3. Life support services represent the services necessary for life in the environment, human communities, and the workplace. They include oxygen for workers in the economy and carbon dioxide that is "breathed" by farmer's fields. Life support services provided by the environment are hampered by growth in the production of economic goods. "Pearce's key point is that the life support system cannot be . . . priced or otherwise allocated by the economy" [Swaney 1987, p. 337].

4. The sink function refers to the fact that all "wastes" from all parts of the environment and from the economy are disposed of in the environment. This sink function can no longer be taken for granted, because overloading the sinks with wastes and pollution from households and production centers increasingly interferes with the environment's other three functions.

CVM and TCM are based on the idea of closed equilibrium systems, and therefore are inconsistent with open systems. The equilibrium concept that defines them is not open to inputs from or outputs to the environment, and therefore CVM and TCM are not modeled to account for those inputs and outputs. "A part viewed in isolation cannot be understood as well, than when viewed (1) in its environmental setting and (2) under consideration of essential interdependencies with other parts" [Mattessich 1978, p. 323].
Nonisomorphic

Real world systems are not isomorphic from part to whole. In isomorphic systems the whole is a reflection of the parts—for example, the sum of the parts. The idea that systems can be studied by looking at individual parts is referred to as reductionism. In living systems, the parts work according to the structure of the system. Work procedures are guided by the requirements of the technology and human consumption is guided by social requirements. GSA allows investigators to accomplish two procedures very important to an investigation. First, it allows for abstracting the system of interest from the overwhelming complexity of the real world. Second, it provides a means of disaggregating the system into subsystems without practicing reductionism. As Rosen has explained, a reductionist hypothesis cannot be true for many of the defined properties of greatest interest about systems. [Rosen 1972, p. 55]. The task thus is to disaggregate or fractionate a system into nonisomorphic systems so that “(a) each of the fractions, in isolation, is capable of being completely understood, and most important, that (b) any property of the original system can be reconstructed from the relevant properties of the fractional subsystems” [Rosen 1972, p. 53]. In this way, subsystems can be effectively used to give us information about the original system.

Real-world systems are not isomorphic reflections from part to whole. Central to GSA is the “notion that a system is characterized by the fact that it is more than the sum of the parts” [Mattessich 1978, p. 20]. Yet the underlying assumption, as well as the operation of CVM and TCM, is that the whole is the sum of the parts. Thus, instead of disaggregating the individual’s beliefs, attitudes, and tastes from the system under study, the reductionist approach is to attempt to sum up the value of natural resources, for example, from the survey findings of individuals. This concept is the reason CVM studies ignore the underpinnings of their own paradigm by assuming that all individuals are equal to the mean user. The investigators want to sum to the whole from the parts.

Equifinality

The equifinality property of systems means that open systems “can reach the same final state from differing initial conditions and along a variety of paths” [De Greene 1973, p. 37]. Because systems are not automatic equilibrium systems, they respond to changes in the external environment to achieve a system goal. Only by adjusting the system
can open systems attain a steady state. The degree of equifinality is reduced as more control mechanisms are introduced [Katz and Kahn 1976, p. 100]. If a technology rigidly sets the requirements of the social system, the flexibility of the social system in dealing with pollution, for example, is reduced.

The concept of equifinality becomes important when determining the restoration of an ecosystem. Since there are alternative paths to achieving system viability, some paths may be less expensive in resources than other paths.

The CV and TC methods do not offer more than one path to explain system value, nor do they offer a way to elucidate the alternative paths within a system.

**System Components**

Real world systems studies, whether they are called sociotechnical, socioenvironmental, or socioeconomic, are concerned with the integration of the components of the social, technical, and natural environmental subsystems. The components of these systems are: (1) cultural values, (2) social beliefs, (3) personal attitudes, (4) technology, (5) social institutions, and (6) the natural environment [Hayden 1982 and 1988].

The CVM and TCM do not attempt to define and deal with the various components of a system. In addition, the way the CVM survey instrument deals with beliefs, attitudes, and tastes makes it impossible, as is described by some of the CVM investigators, to know what the results of the survey mean. In the first place, this survey instrument is attempting to measure a taste and tastes are not important for valuing social or ecological systems. In addition, as explained by investigators who complete CV surveys, beliefs overwhelm tastes. This is consistent with findings in social psychology, because beliefs and systems of beliefs (ideology) are the basic social criteria, and the determinants of attitudes and tastes. This means, according to Daniel Kahneman,

that we should exercise great caution in measuring option values and reservation values, because responses that are obtained in such measurements are likely to be heavily loaded with ideological content. . . . The key observation is that there is a class of problems in which people’s answers to preference questions seem quite insensitive to the numbers that are mentioned in these questions. Indeed, people seem to be ready with an answer before the relevant numbers are specified [Kahneman 1986, p. 190].

Kahneman points out that people have their minds made up on what
they want done; for example, if they want the environment cleaned up, no matter what CV questions are asked or how the questions are specified, the answer is the same—it reflects (but does not measure) the beliefs at which the respondents have arrived. That is why demand functions for very different cleanup operations come out strikingly similar. The respondents are not considering the monetary price. In a Canadian CVM study, "the results indicate that people seem to be willing to pay about as much to clean up one region or any other, and almost as much for any one region as for all Ontario together" [Kahneman 1986, p. 191]. The failure to distinguish among beliefs and tastes in a system context prevents the CV approach from obtaining relevant data.

Control and Regulation

Crucial to systems and therefore an important focus of GSA are the control and regulation mechanisms of systems. System control and regulation takes place through rules, requirements, and criteria. Two types of control are emphasized in GSA.

The first type of control is that every system element or subsystem that makes a delivery to another element or system exerts control "if its behavior is either necessary or sufficient for subsequent behavior of the other element or system (or itself), and the subsequent behavior is necessary or sufficient for the attainment of one or more of its goals" [Ackoff 1971, p. 670]. This is a control through relationship and requirement linkages. An example is the effect of habitat cover on the kind and structure of wildlife in the habitat.

However, before elements or systems can perform the behavior pattern that gives them the first type of linkage control, other control mechanisms and rules are needed to determine their behavior. These constitute the second type of control. "Biological and social structures are not objective in the sense of physical laws. They are coherent systems obeying dynamic laws and syntactical rules that are distinguished from isolate physical systems by their ability to change their internal constraints and thereby change the rules of the game" [Pattee 1976, p. 179].

DNA is an example of system rules that give DNA extraordinary authority over the cellular collectivity, and the development of multicellular organisms . . . shows that the cells do not simply aggregate to form the individual, as atoms aggregate to form crystals. There are chemical messages from the collections of cells that constrain the detailed genetic expression of individual cells that make up the collection. Although each cell began as an autonomous "typical" unit with
its own rules of replication and growth, in the collection each cell finds additional selective rules imposed on it by the collection which causes the differentiation [Pattee 1973, p. 77].

The presence of controls and constraints in a system is a distinguishing characteristic of living systems.

Technology is another example of system rules. It provides requirements for social systems. These are often in the form of criteria that must be met. The technical component "contributes preeminently to the self-regulating features of the system" [De Greene 1973, p. 47]. "Thus the technological system sets requirements" [De Greene 1973, p. 47].

In social systems, primary rules are social belief criteria. They give the social system structure. "Social structure consists of myths, constraints, rules, customs, beliefs, legal codes, and the like. These structure social systems by guiding social and economic action, by legitimizing transactions, and by requiring delivery to be made" [Hayden 1986, p. 386]. As James Swaney has clarified, in addition to the cellular, technological, and social, there are ecological constraints, rules, and criteria that we are attempting to ignore in modern real world systems [Swaney 1985]. These are also part of the system, and an attempt to override them will degrade the system.

The market approaches to natural resource valuation do not include any explanations of the control and regulation mechanisms in the system that direct the socioecological system. The CV and TC survey questionnaires separate individuals from the system and turn the natural environment into an isolated object, and then request that individuals respond to objects without respect to eliciting or explaining social, technological, or ecological criteria or control requirements.

**Hierarchy**

Following the discussion on system control devices, it is probably not surprising that all systems experience hierarchical arrangements of many kinds. Laszlo has defined hierarchies as "higher order systems which within their particular environments constitute systems of still more indecisive order" [Laszlo 1972, p. 19]. Pattee emphasized the control aspects of hierarchy. He wrote:

In a control hierarchy the upper level exerts a specific dynamic constraint on the details of the motion at a lower level, so that the fast dynamics of the lower level cannot simply be averaged out. The collection of subunits that forms the upper level in a structural hierarchy now also acts as a con-
constraint on the motions of selected individual subunits. This amounts to a feedback path between levels. Therefore the physical behavior of a control hierarchy must take into account at least two levels at a time [Pattee 1973, p. 77, emphasis added].

The emphasis was added to Pattee’s quote to emphasize criteria methodologies need to meet. CVM and TCM do not meet these criteria.

There is no attempt with CVM or TCM to define or determine the system hierarchy or to determine the relationships among different levels in the hierarchy. There is an assumption that the utilitarian moral principle of maximizing individual utility is a criterion that should be placed above all other criteria in a system [Rohrlich 1976, p. xxiii].

**Flows, Deliveries, and Sequences**

Systems could be defined as flows of sequenced deliveries. The concept of flow is fundamental to systems.

Internal and external descriptions of systems are wholly complementary approaches to modeling systems structures and this equivalence can be seen through the unifying concept of flow. If a system has been described internally in terms of a number of state variables between which are defined certain relational functions, then these state variables can be considered to change as a result of flows occurring [Bryant 1980, p. 73].

Through input flows from the natural system into socioeconomic systems, resource analysis is completed. It is also important to explicitly include output flow to determine environmental impact assessment and valuation. “The delivery flow through the process is the substance of socioeconomic life, and is a way to measure thresholds of change. Within a system, there are tolerance levels with regard to variation of deliveries” [Hayden 1986, p. 387].

Systems respond to flows according to the level, or amount, of the flow. It is through flow levels that systems are integrated. For example, the level of aggregate demand delivered in the economy influences the level of employment. Delivery levels outside the tolerable threshold will create negative feedback for change. For example, the food deliveries may be inadequate or the air pollution level too great.

The goal of CVM and TCM is to measure the monetary value of the flow of utility to individuals. There is no attempt, however, to define system flows or to outline the network of deliveries and sequences.

**Negative and Positive Feedback**

For policy purposes, especially with regard to the natural environ-
ment, the system concept of negative and positive feedback is very important. “Negative feedback is associated with self-regulation and goal-direction, positive feedback with growth and decay” [De Greene 1973, p. 22]. The inputs of living systems consist not only of energy and material, but also of information, all of which “furnish signals to the structure about the environment and about its own functioning in relation to the environment” [Katz and Kahn 1966, p. 95]. Feedback is a form of inter- and intra-systemic communication in which the past performance of the system yields information to guide its present and future performance. Negative feedback systems are error-activated and goal seeking in that the goal state is compared with information inputs on the actual state and any difference (error) provides an input to direct the system toward the goal state. Negative feedback thus leads to the convergence of system behavior toward some goal. “When the system’s negative feedback stops, its steady state disappears, and the system terminates” [De Greene 1973, p. 78]. It has been argued that one of the main benefits of democracy is the negative feedback and interference from the citizen’s government that serves as the comparator to evaluate the condition of the system.

What makes the open systems approach so vibrant from a policy standpoint is the fact that it views the environment as being an integral part of the functioning of a sociotechnical system. Thus, external forces that affect the system need to be included in the system. Furthermore, negative feedback mechanisms are needed to provide information about environmental changes that will affect the system, in order to better understand what, if any, policies need to be made to insure a continued effective system.

Positive feedback systems, in which positive feedback information overwhelms negative feedback information, tend to be unstable if a change in the original level of the system provides an input for further change in the same direction. “Society and technology tend to reinforce one another in a positive feedback manner, which is not always desirable. At the same time there is often a loss of negative feedback and self regulation” [De Greene 1973, p. 7]. For example, if an agricultural system based on advanced technology is not incorporating the negative information regarding soil erosion, the system will continue its growth until destruction.

With the CVM and TCM techniques, respondents to the surveys are not allowed to have negative or positive feedback information from the rest of the system if their responses are used for making policy. They are not presented with alternative system consequences; thus, they are not allowed to make error-activated responses, as they would in a sur-
vey that attempted to replicate responses in a democratic system.

Even within the CVM context, as Smith and Desvousges point out, results are influenced by the failure to generate feedback from which respondents can learn. "An important source of the available empirical evidence, and laboratory experiments suggest that individuals may have difficulty in dealing with the concept of compensation. This is especially true when there is no opportunity for individuals to learn about transactions that involve compensation through experience" [Smith and Desvousges 1986, p. 291].

*Differentiation and Elaboration*

"The unique character of biological and social system behavior that distinguishes them from non-living systems is their tendency to evolve greater and more significant complexity" [Pattee 1978, p. 99]. This idea has been expressed in almost all disciplines. Katz and Kahn have stated with regard to social systems that "open systems move in the direction of differentiation and elaboration. . . . Social organizations move toward the multiplication and elaboration of roles with greater specialization of function" [Katz and Kahn 1976, p. 99]. David Hunter and Phillip Whitten explain a similar evolution with regard to the economy. "In the economic sphere, a traditional society displays relatively little division of labor, but modern societies produce a proliferation of highly differentiated and specialized occupational statuses and roles" [Hunter and Whitten 1976, p. 287]. Differentiation becomes an important characteristic when discussing ecological restoration. It is important to think about future differentiation potential when considering option values of an ecosystem.

The CV and TC methods, consistent with their neoclassical base, do not deal with system differentiation and elaboration.

*Real Time*

The time concept most consistent with GSA is system real time, which is inconsistent with classical ideas about time. According to the classical Kantian system, for example, there are the so-called forms of intuition, space and time, and the categories of the intellect, such as substance, causality and others which are universally committed for any rational being. Accordingly science based upon these categories, is equally universal. . . . Newtonian time and strict deterministic causality, is essentially classical mechanics which, therefore, is the absolute system of knowledge, applying to any phenomenon as well
It is a well-known fact that modern science has long recognized that this is not so [Von Bertalanffy 1969, p. 226].

Modern science applies the time concept most appropriate for the subject under investigation. “The biologist finds that there is no absolute space or time but that they depend on the organization of the perceiving organism” [Von Bertalanffy 1969, p. 229]. A similar idea is found in the concept of experienced time. “Experienced time is not Newtonian. Far from flowing uniformly . . . it depends on physiological conditions” [Von Bertalanffy 1969, p. 236].

Time is not a natural phenomenon; rather, it is a societal construct. The construct should be consistent with the GSA view and counter to the reductionist view. Time, if it is to be a useful tool in, for example, ecological restoration, should be what usually is connoted by the word timeliness. Timeliness requires that we ask the question: which restoration project will sequence and deliver the right amount of system components and elements at the right points in the ecosystem and sociotechnical system to allow for integration, maintenance, and restoration? “Temporal evaluation that judges whether a project correctly sequences the delivery of impacts with system needs is consistent with the basic concepts of computer science real time. Real time systems relate to the sequential events in a system, rather than to clock or calendar time. The system itself defines when events should happen” [Hayden 1988, p. 346].

Real time is not used in CVM and TCM studies.

Evaluation and Valuation

Methodologies should be evaluated for their contributions to the solution of the overall problem, to embrace the subtleties of the value of wildlife, to apply a broad definition of ecosystems, and to provide information regarding the value of habitats to take full account of regulations and policies on the environment. That approach to ecosystem evaluation and valuation is consistent with the GSA context. As A.D. Hall and R.E. Fagen have stated, “analysis, evaluation and synthesis of systems is not concerned primarily with the pieces . . . but with the concept of system as a whole; its internal relations, and its behavior in the given environment” [Hall and Fagen 1968, p. 92]. The focus of evaluating and valuing is to identify the value of the various entities as they contribute toward making the socioecology viable. (See Mattessich [1978] and Laszlo [1972]). Viability includes the idea that there be redundancy in the system network and deliveries to maintain system
sufficiency. Valuation assists in making decisions about the maintenance, coordination, and restoration of systems through the coordination and sequencing of relevant events.

The market approaches, CVM and TCM, as stated above, are concerned with evaluating prices separate from the system. This is inconsistent with GSA evaluation of the various entities as they contribute toward making the socioecology viable. The measurement of people's contingent market value of the environment implies that nature's only purpose is for the enjoyment of human beings. When an attempt is made to apply market demand functions for non-market goods, this implies that the environment has value only for humans. These valuation techniques ignore that ecosystems, or particular flora and fauna, have functions other than human demand for them. Thus, the CV and TC techniques are inappropriate for establishing a systems evaluation or cardinal ranking of nonmarket goods.

The CVM treats individuals as the abstract Economic Man explained by Majumdar.

The Economic Man is truly the knight of popular mythology. His is the solitary figure of the Subject facing the Object, which is the rest of the universe. In this Subject-Object relationship the Economic Man has no collaborator and no human opponent. . . . What requires further emphasizing is that his motives are constructed to be purely monetary. From which two attributes of the Economic Man clearly stand out. In the first place, he is unaffected by (and incapable of affecting) what happened to others. In the second place, he would not pursue a target which could not directly or indirectly be brought into relationship with the measuring rod of money [Majumdar 1975, p. 3].

This definition of economic man is classical economic man rather than neoclassical person. The first point is that the way CVM is operationalized, neoclassical person is forced into the mold of the classical subject-object (respondent-questionnaire) economic man. The second point is that it is not necessary to take either the classical or neoclassical utility approach to valuation. “In fact, a dominant section of contemporary political and social philosophy appears to be built specifically upon its denial” [Majumdar 1975, p. xiii].

Conclusion

Although the CVM and TCM are used extensively for measuring the value of ecosystems, neither method can be legitimized in a theoretical or applicable sense from a neoclassical, psychometric or general sys-
tems point of view. The CV and TC approaches lack methodological, theoretical and empirical grounding. Their continued use will mislead valuation attempts and frustrate policy intended to restore a viable environment.

References


