Final Report

Introduction to Attribute-Based Stated Choice Methods

by Wiktor Adamowicz, Jordan Louviere, Joffre Swait

January 1998

Submitted to

Resource Valuation Branch
Damage Assessment Center
NOAA - National Oceanic and
Atmospheric Administration
US Department of Commerce

Pursuant to NOAA Purchase Order 43AANC601388

January 1998

Final Report

Introduction to Attribute-Based Stated Choice Methods

by Wiktor Adamowicz, Jordan Louviere, Joffre Swait

Submitted by

ADVANIS
Suite 1600, Sun Life Place
10123 - 99 Street
Edmonton, Alberta
Canada T5J 3H1
Phone (403)944-9212 Fax (403)426-7622

TABLE OF CONTENTS

1. Introduction

Several major environmental statutes enacted in the 1970s designate resource management agencies as trustees of the natural resources on behalf of the public, and enable the trustees to recover damages for injuries to public resources from releases of hazardous substances and discharges of oil. The standard measure of damages in the various statutes is the cost of restoring the resources to baseline conditions ("primary restoration") plus compensation for the interim loss of resources from the time of the incident until full recovery from the injuries.[1] Though the measure of damages for interim losses was originally characterized in terms of *monetary compensation*,[2] trustees are only allowed to spend their damage recoveries on enhancing or creating ("restoring, rehabilitating, replacing or acquiring the equivalent of") natural resources.

The statutory restriction on the use of the recoveries has motivated the development of an alternative *resource compensation* measure of damages for interim losses -- the cost of "compensatory restoration" actions that provide sufficient resources and services to compensate for the losses. The National Oceanic and Atmospheric Administration (NOAA) incorporated the resource compensation measure in its 1996 regulations implementing the natural resource liability provisions of OPA.[3] Under the new rule, the focus of the analysis becomes identifying and evaluating the attributes of environmental resources (and the services they provide) at both the injury and replacement resource sites.

Current methods of environmental valuation can be categorized as Revealed Preference methods (RP) and Stated Preference methods (SP). RP methods (such as travel cost or hedonic price models) can be used to construct resource compensation profiles to a certain extent. For example, random utility models based on RP information can be used to determine how attributes from some alternatives in the choice set can be used to compensate for damages to attributes in another alternative. However, RP data can be found wanting for a number of different reasons subsequently reviewed, not least of which is that the behavior of interest may not be observable. Furthermore, RP data may suffer from a variety of problems

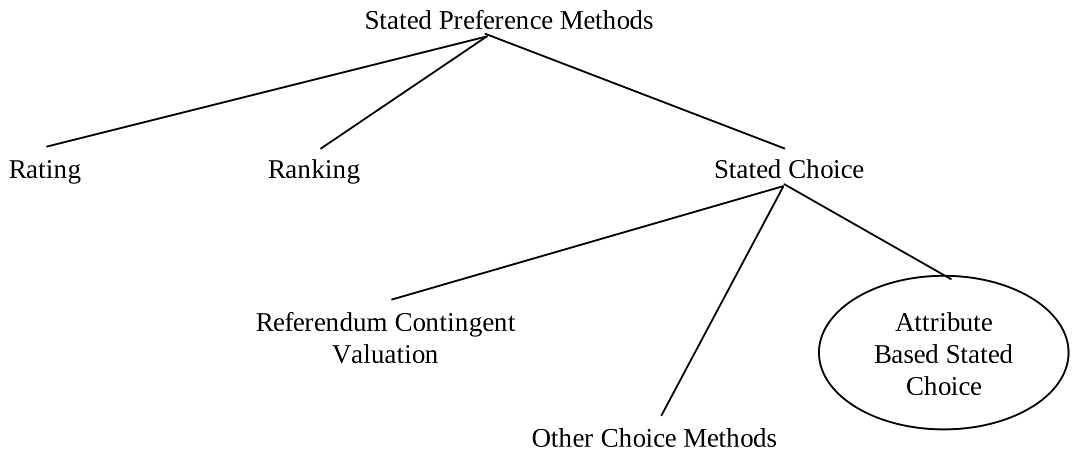
---

[1] Throughout this discussion we assume that compensation is determined by evaluating the appropriate individual measures of welfare (compensating or equivalent variation). Clearly, this is a specific normative approach that is not without controversy. In practice this approach almost always ignores interactions between agents or interdependent preferences which may be very important in the case of public goods (Heckman, 1997). Further-more, the approach assumes citizen or consumer sovereignty rather than rely on elected representatives, through some preference revelation mechanism, to place value on the public goods. Nevertheless, this approach is consistent with previous approaches to the valuation of losses.

[2] Actually, valuation methods like contingent valuation are typically used to develop estimates of willingness-to-pay for *prevention* of environmental damage rather than willingness-to-accept compensation for the damage, as the latter is more difficult to elicit. For details on this issue see Mitchell and Carson (1989)..

[3] 15 CFR §990. The US Congress is currently considering bills to reauthorize CERCLA that would incorporate these concepts into statutory language. Absent statutory reauthorization, the US Department of Interior has indicated it plans to incorporate these concepts into a revised rule, during its ongoing biennial review of its natural resource damage assessment regulations implementing CERCLA.

Figure 1: Stated Preference Methods



that limit their usefulness in model development and compensation determination (see Adamowicz et al. 1994).

Figure 1 outlines the variety of SP methods that have been employed in marketing, transportation, economics and other literatures. This figure illustrates that referendum contingent valuation is one form of stated choice method based on random utility theory. "Open-ended" contingent valuation falls into the category of "ranking" in the figure above since it essentially involves ranking a scenario with a monetary metric (versus some other metric, such as a rating scale). In this paper we focus on "Attribute Based Stated Choice Methods," since these methods have not been employed in environmental valuation to a great degree and they may be of significant value in resource compensation cases. However, it should be noted that if price or "tax" is the only attribute being considered in an Attribute Based Stated Choice Task, then the task essentially collapses to a referendum contingent valuation situation. In the remainder of this paper when we refer to Stated Choice Methods we are referring to Attribute Based Stated Choice Methods, recognizing that referendum contingent valuation can be considered a subset of this approach. Furthermore, this reveals that many of the issues surrounding referendum contingent valuation also affect Stated Choice Methods. However, we will focus on outlining Stated Choice Methods and concentrate on differences between these methods and traditional contingent valuation methods.

As mentioned above, the most commonly used SP method in environmental valuation is contingent valuation (Carson et al. 1994b). While this technique provides a method for determining the monetary value of environmental damages, it may not be well suited to the determination of sets of attributes that could be used for compensation.[4] In order to develop

[4] It is worth noting that there have been applications of contingent valuation that have attempted to elicit attribute type responses. Boyle, Welsh and Bishop (1993), for example, employ contingent valuation in such a fashion. While no formal test between approaches has been done to our knowledge, we suspect that stated choice methods designed to elicit attribute based responses will be better at providing such information.

such compensation profiles one may wish to employ stated choice techniques that have been developed in the marketing and transportation literatures to evaluate attributes. These methods are subsequently described in more detail, including their advantages and disadvantages. Finally, SP and RP methods can be combined to overcome some of the difficulties associated with each approach (see Adamowicz et al. 1994, Cameron 1992, and Swait, Louviere and Williams 1994). This combination approach is particularly useful in cases where observed behavior information is available.

While, conceptually, both SP and RP methods can be used to determine resource compensation, there has been little experience in using either technique for such purposes. In this paper we examine the use of SP methods, particularly stated choice, in environmental valuation and in determining resource compensation.

The structure of the paper is as follows: we next present a conceptual model of resource based compensation in Section 2, which lays the groundwork for the overview of stated choice methods in the Section 3.

2. A Conceptual Model of Resource Based Compensation

2.1 Ecological conditions

In attempting to determine the appropriate compensation for a case of environmental damage,[5] the set of natural resource services that may be considered as compensatory could be very large. For example, injury to a particular fishing site may be compensated by enhancing the existing site (thus increasing future flows of benefits), enhancing an adjacent site, creating new sites, improving ecological services "upstream" that will potentially improve several sites, and/or enhancing ecological or human services at sites that are not related to the injury site. Before attempting to assess social preferences for these various options it will be necessary to understand, at least to a certain degree, the biological relationships in these systems.

The biological relationships are required to understand the spatial and temporal linkages between the restoration action and the result. For example, if fish habitat improvements are proposed as a method of resource compensation, biological relationships between the habitat program and the spatial and temporal distribution of fish will need to be understood and presented to individuals in a fashion that they can understand. Often these relationships are uncertain and it may be necessary to convey some notion of the degree of uncertainty of the biological result to individuals when they are evaluating the options.[6] The key consideration here is that in resource-based compensation the understanding of biological relationships, and

[5] The context used throughout this section is a case of compensation for environmental injuries. However, measurements of resource based valuation for environmental gains could also be addressed using the same framework.

[6] See Cameron (1997) for an approach for incorporating respondent uncertainties into stated preference measures of welfare.

the presentation of these relationships to individuals, takes on a larger role since the scope of regions, species and habitats to be valued may be substantially increased, relative to more focused monetary valuation efforts.

The effective presentation of biological information will require close contact between biological and social scientists and may require an iterative approach to determining appropriate compensation. For example, an exercise can be developed in which the trustees construct potential resource compensation packages from a broad set of potential options. These options can then be submitted to biology and ecology experts to develop more detail on potential outcomes and information gaps. The resulting packages will then be returned to the trustees for further consideration in light of the new information.

The ecological conditions, by and large, form the supply or technical relationships required for the determination of resource-based compensation measures. Understanding the supply relationships is necessary for the development of accurate assessment tools. If ecological relationships are misinterpreted or misunderstood, the resulting analysis of human preferences will be flawed. This is not to say that complete knowledge of the ecological conditions is required before human preference assessment can begin. Nonetheless, preference assessment methods may need to incorporate uncertainty and probabilistic relationships (Erdem 1993; Finn and Louviere 1992; Gregory et al. 1993; Hayes et al. 1995; Keeney and Raifa 1976; McDaniels et al. 1995).

2.2 Human Preferences

In general, individuals may be said to gain utility from the natural environment through two pathways: ***direct uses*** (recreation, commercial harvesting, etc.) and ***passive uses***[7] (preferences over habitats, ecological services, etc.). These pathways are related because preferences over habitat, through ecological relationships, may also affect direct use values. Thus, it is mostly for purposes of discussion that we keep these pathways separate. In attempting to develop measures of natural resource compensation, one needs to examine preferences over a wide variety of attributes and uses and understand the relationships among these various elements.

It is necessary to recognize use values in the broadest sense. Use values are typically associated with direct uses of the resource (recreation, etc.). However, indirect use values are also very important. Indirect use values include the impact of ecological services on water quality, nutrient flows, air quality, erosion control and other aspects of the environment that are linked to direct use activities like recreation values and property values. These indirect use values (or “off-site” direct use values) provide a fuller picture of the linkage between human uses and environmental service flows.

[7] Passive use values are typically distinguished from use values in that the former have no easily traced observable behavioral trail. This distinction arises in our discussion of the utility arising from these different types of environmental benefits. Ecological service flow changes may affect both use and passive use values.

Utility theory provides the conceptual basis for resource compensation measures since all attributes of ecological services can be represented as entering an individual's utility function. For example, ecological services could enter an indirect utility function as a vector of attributes (q) along with prices (p) and income (M). A measure of resource compensation (RC) for a change from initial ecological attribute level q_0 to a lower (after "damage") level q_1 will be represented as:

$$V(p, M, q_0) = V(p, M, q_1 + RC) \quad , \tag{1}$$

where RC represents the changes in ecological attributes required to make utility in the "injury" case equal to utility before the injury. Equation (1) represents the elemental form of resource compensation since the entire vector of ecological service attributes is assumed to be represented in q and these attributes can be represented in a single indirect utility relationship (see Jones and Pease 1997, Desvousges et al. 1997 and Adamowicz and Swait 1997 for further discussion of the utility basis for resource compensation). What is more likely is that there is an utility tree/function[8] of products/services (corresponding to components of a separable utility function) that are affected by subsets of environmental attributes.

As an example, consider an individual who gains utility from recreational fishing (one sub-branch of the utility tree) and also gains utility from the knowledge that certain ecosystems exist and are functioning (another sub-branch). This individual may have alternative sites to choose for recreational fishing and these sites also provide "habitat" utility for the individual. Some sites, however, do not provide utility as recreation sites. As illustrated in Figure 2, the individual gains utility from three recreation sites, while all five "sites" provide habitat from which the individual derives utility. In the figure, there is overlap between the sites. Thus, changes in ecological conditions at one of the sites may affect the attributes of the recreation experience. Also, there will likely be linkages between the sites in that changing ecological conditions at one site, via these linkages, may affect the quality of the recreational experiences at another site.

This representation of preferences from direct uses (recreation) and habitat preferences (passive use) is employed to illustrate the fact that knowledge is required about the attributes of the elemental alternatives (recreational fishing sites, ecological service regions) as well as the preference relationships between these sites (the utility tree). Knowledge of the effect of attributes on any single set of alternatives will not be enough to determine resource compensation; knowledge of the utility tree structure (specifying the degree of separability between classes of goods and services) will also be required. In addition, as described above, knowledge of the ecological relationships between elemental alternatives will be necessary for a complete understanding of the potential for resource compensation. Furthermore, in order to maintain consistency with economic axioms of choice, linkages with other categories of goods and services and the presence of a budget constraint will be necessary to elicit economically meaningful resource compensation measures.

---

[8] For example, a utility tree can be constructed with an upper level for aggregate goods (food, clothing, etc.) and lower levels for disaggregated goods (meats, dairy, vegetables, etc.)

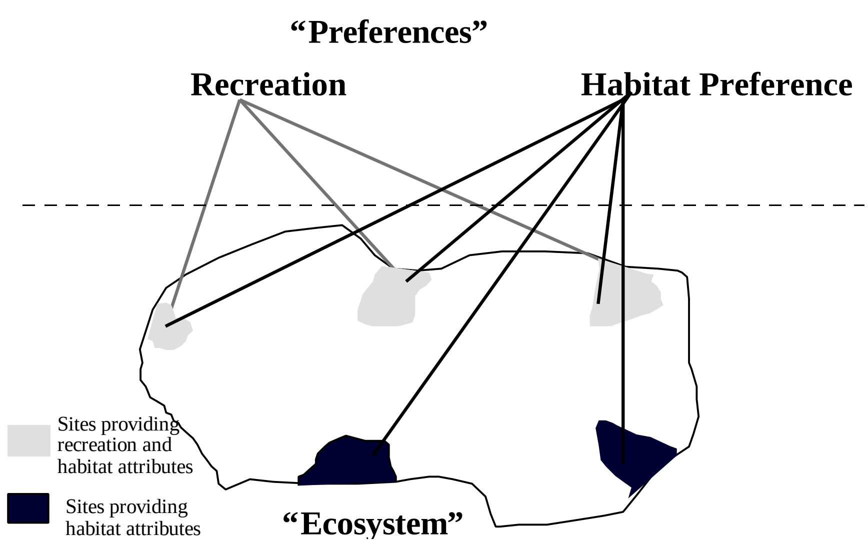
Figure 2 illustrates that the determination of resource compensation will require an understanding of the preference structure of individuals in terms of the branches of their utility functions (trees), the attributes of relevance within these branches and the ecological conditions that link these attributes. Injury to a recreation site can be compensated with improvements to another recreation site, or to some of the ecological services that affect habitat preference. Understanding of preferences is required to know which alternatives should be improved, and to what extent. In addition, an understanding of the regulatory and institutional bounds associated with resource compensation is needed. For a variety of reasons (e.g. institutional, economic, equity) unrelated to ecological factors or individual preferences, certain forms of resource compensation will not be deemed appropriate. For example, creation of new recreation sites a great distance away from the injury site may be deemed inappropriate on equity grounds; institutional constraints may also place bounds on the resource compensation measures.

We now turn to a description of Stated Choice Methods and how they may be employed in resource based compensation.

3. Stated Choice Methods

The term Stated Choice Methods (SCM), as we use it in this paper, refers to a flexible approach to collecting preference data (generally, choices and rankings, whether full or

Figure 2: Preferences over Environmental Attributes



partial) from subjects in hypothetical situations.[9] While this is generally done with paper-and-pencil tasks, the elicitation scenario can become quite elaborate, involving videotapes, computer simulations, virtual reality, etc. The objective, however, is to place the decision-maker in a realistic frame of mind to compare a number of alternatives, each described in terms of some number of attributes. The decision context and product descriptions are the *stimuli*, and the individual's decision (which may be a choice, a ranking, or a quantity) is the *elicited response*. The decision scenario and product descriptions are most commonly generated using experimental design techniques, with the objective of minimizing the number of combinations that must be given to respondents to enable statistical identification of the underlying preference functions. It is common practice to have decision makers view multiple scenarios, motivated by a desire to utilize resources efficiently.

Choice and ranking data generated from SCM are generally analyzed using Random Utility theory and utility maximization as a conceptual framework. Thus, the same econometric methods used to analyze Revealed Preference (RP) data are employed with choice and ranking data from SCM. With quantity responses (e.g. number of trips, days at camp site), appropriate analysis techniques also already exist: Ordinary Least Squares, Tobit models, count models, etc.

The advantages of using SCM are manifold: (1) control of the stimuli is in the experimenter's hand, as opposed to the low level of control generally afforded by observing the real market place; (2) control of the design matrix yields greater statistical efficiency and eliminates collinearity (unless explicitly built into the design); (3) more robust models obtain because wider attribute ranges can be applied than are found in real markets; (4) introduction and/or removal of products and services is straightforwardly accomplished, as is the introduction of new attributes. The last point, in fact, is often practically impossible, but certainly always difficult, in actual markets.

Thus, SCM are not a theory of behavior; rather, they are simply a means to *generate behavioral data* from consumers. Well-established consumer choice theories and econometric modeling techniques can be applied to such data, just as they are applied to RP data.

3.1 Historical Antecedents of Stated Choice Methods

It is often suggested that Stated Choice Methods that focused on attributes (hereafter, SCM) evolved out of the conjoint analysis paradigm long associated with marketing research.[10] However, early applications of SCMs were developed by Louviere and Hensher (1982) and Louviere and Woodworth (1983) as a natural analog to already well-established revealed preference choice modeling theory and methods. While it is true that conjoint analysis is a very popular approach for understanding and predicting consumer tradeoffs and choices, conjoint

[9] As described in the introduction, referendum (and discrete choice) contingent valuation can be considered a form of stated choice methods.
[10] An overview of conjoint analysis can be found in Louviere (1988a). Please see the Glossary for a fuller description of conjoint analysis.

analysis is not, nor can it be, a theory of choice behavior.

Rather, conjoint analysis is a theory about the behavior of numbers in response to systematic manipulation of levels of some set of attributes, or more generally, independent variables. The reason we say this is because traditional, ratings-based conjoint analysis presents the decision maker with a single product description, then elicits a response (e.g. likelihood to purchase, attractiveness of product) on a rating scale (for example, 1=highly unlikely to buy, …, 7=highly likely to buy). If the numbers observed in a conjoint analysis are valid indicators of preference produced by individuals, and the numbers satisfy a number of mathematical axioms, then such numbers can be used to make inferences about consumer preferences. We are unaware of any formal theory, however, which allows one to use numbers produced for or derived from conjoint analysis to predict choices. In fact, conjoint analysis, as originally conceived in psychology and as practiced by conjoint analysts in marketing, makes no attempt to establish a direct connection with economic theory, especially in the case of measuring or modeling consumer surplus or other values of policy interest to natural resource economists. Thus, conjoint analysis (and for that matter, any stated preference method) needs a strong behavioral theoretical foundation consistent with economics. Lacking such a link, methods like conjoint analysis are atheoretical or non-behavioral.[11]

We now turn our attention to the conceptual foundations of SCM, a method of preference elicitation that differs from conjoint analysis because it adopts economic theory as its backdrop. Special emphasis is placed on SCM's origins in Random Utility theory.

3.2 Behavioral Foundations of SCM

SCM behavioral foundations include

1. <u>Lancastrian consumer theory</u> (Lancaster 1966), which proposes that utilities for goods can be decomposed into separable utilities for their characteristics or attributes;

2. <u>Information processing</u> in judgment and decision making in psychology (e.g., Hammond 1955; Slovic and Lichtenstein 1971; Anderson 1970, 1981, 1982); and

3. <u>Random utility theory</u>, which forms the basis of several models and theories of consumer judgment and decision making in psychology and economics (e.g., Thurstone 1927; McFadden 1974; Manski 1977; Yellott 1977).

[11] Another objection to conjoint analysis is the statistical method usually employed, which is OLS. The implicit assumption in such models is that the underlying latent variable is interval scaled, which has been repeatedly shown not to hold in practice. Other methods, such as ordered probit, can be (but have rarely been) employed to analyze the ratings appropriately. Even so, however, there remains the objection that there is no theoretical link between the measured ratings and the ultimate behavior of interest (most often choice).

Individuals or decision makers come to recognize a need to solve problems, make choices, or obtain benefits, which initiates search and learning to find out what solutions are available to meet needs. Environmental conditions and human actions impact individual perceptions of the positions that various alternatives occupy on a set of key decision dimensions (attributes) on which individuals base their evaluations and comparisons. Note that these dimensions of decision maker evaluation need not correspond to the “engineering” view of the attributes; instead, they may constitute intermediate constructs that include one or more physical attributes. For example, a recreational fishing site may be evaluated in terms of its “fishing quality,” which is captured by attributes such as “% of fish caught that are over a certain size,” “fish caught per hour,” “size of fish caught,” etc. Lancaster’s (1966) approach allows for this information integration process. Note also that the alternatives could be items that are very similar to products in a market, like recreation sites, or they could be options in a referendum or public choice structure.

Positions occupied by alternatives on key attributes are evaluated by individuals relative to "how good" it is to have such and such a level of an attribute (e.g. quality) compared to the level possessed by competing alternatives. As well, individuals tradeoff levels of one attribute against levels of other attributes, implicitly weighing and valuing both the attributes and the positions occupied. The latter tradeoffs lead to the formation of preferences for various alternatives, ultimately resulting in the choice of which particular option to choose, be it a recreational fishing site or an alternative in a referendum.

Given descriptions of alternatives on key attributes, SCM allow one to understand and model how individuals evaluate product attributes and choose among competing offerings.

3.3 Random Utility Theory and Choice Model Specification

Thurstone (1927) proposed random utility theory as the basis for explaining dominance judgments among pairs of offerings. As conceived by Thurstone, consumers should try to choose the offerings they like best, subject to constraints (e.g., income, time), just as in standard economic theory. However, a consumer may not choose what seems to the analyst to be the preferred alternative. Such variations in choice can be explained by proposing a random element as a component of the consumer's utility function. That is,

$$U_i = V_i + \varepsilon_i, \tag{2}$$

where U_i is the unobservable, true utility of offering i; V_i is the systematic (i.e. known) component of utility; and ε_i is the random component. The econometric justification for this random component is that the analyst may omit variables or commit measurement errors, the consumer may be inattentive to the particular decision, etc.

The presence of this random component permits the analyst to make probabilistic statements about consumers’ behavior. Thus, we focus on modeling the probability that a consumer will choose the i-th offering from some set of competing offerings, say C, which can be expressed as:

$$P(i|C) = \Pr[U_i > U_j] = \Pr[(V_i + \varepsilon_i) > (V_j + \varepsilon_j)],\ \forall\ j \in C. \tag{3}$$

The systematic component of utility is that portion of product attractiveness that can be related to product attributes; our ability to capture it depends on how well we identify, measure and include as many of the key factors that influence choice as possible. Thus, analysts must devote sufficient time and resources in advance of data collection and modeling to identify and include as many of the key influences on choice as possible. This can be based on primary, qualitative research (e.g. focus groups) that is tailored to a particular project, from secondary research (e.g. literature sources, previous experience with the same or similar products), or (as is most common) from a hybrid approach that uses both secondary and primary research. The goals of this phase of any SC model should be (i) to identify how consumers think about the evaluative process (i.e. the evaluation dimensions and their levels, the links between evaluation dimensions and engineering or physical attributes) and (ii) to identify which of those dimensions should be included in which of the two components of the utility function (i.e. the systematic and random components).[12]

Once identified, the analyst has to specify how these variables combine to drive systematic preferences. That is, the analyst must propose a utility function to specify the formal relationship between the explanatory variables and choice behavior. With no loss of generality, the systematic component can be expressed as a linear-in-the-parameters function of the explanatory variables as follows:

$$V_i = \beta' \mathbf{x}_i \tag{4}$$

where β is a k-vector of utility coefficients associated with a vector $\mathbf{x}$ of explanatory variables (including income, prices, other attributes of the alternative and interactions between these elements). Equation (3) can then be rewritten as follows:

$$P(i|C) = P[(\beta' \mathbf{x}_i + \varepsilon_i) > (\beta' \mathbf{x}_j + \varepsilon_j)],\ \forall\ j \in C, \tag{5}$$

where all terms were defined earlier. Equation (5) indicates that the probability that a consumer will choose offering $i \in C$ equals the probability that the combined systematic and error components of offering i are higher than the systematic and associated error components for all other competing offerings. Equation (5) also suggests that our objective is to identify and estimate the β vector associated with the variables hypothesized to explain choice. Choices may differ systematically from individual to individual, and to account for as many of these individual differences as possible, the set of explanatory variables can be expanded to

[12] It is not usual in the literature to view the error component of choice models as an important part of the specification (though see, e.g., Swait and Adamowicz 1997, Hensher, Louviere and Swait 1997, Louviere and Swait 1996). We address this issue later in the text.

include individual difference (i.e. demographic and psychographic) measures **z**, with associated vector of coefficients γ. These individual difference measures may be hypothesized to influence utility levels via intercept and/or slope coefficients in the β vector.

Many different probabilistic choice models can be derived by making different assumptions about the distribution of the errors (random component). For example, a bivariate normal distribution yields the binary probit model (Thurstone 1927), which has its multivariate generalization in the Multinomial Probit discrete choice model; a Gumbel distribution gives rise to the conditional or Multinomial Logit (MNL) model (McFadden 1974; Ben-Akiva and Lerman, 1985); and a Generalized Extreme Value distribution gives rise to models such as the Nested MNL (McFadden 1981) and the Ordered GEV (Small 1981).

The systematic (or mean) component of utility can be identified and the parameters estimated from an appropriately designed empirical study of the way in which choices vary in response to differences in the positions of the offerings on key attributes and differences in individual decision makers. That is, individuals' preferences are revealed by the choices made in choice experiments. "Stated choices" are decisions made in hypothetical markets in which there may be no corresponding real choices, or any "real" consequences of making a choice. Despite the potential lack of realism in such markets, random utility theory nonetheless suggests that consumers should try to maximize utility, although the utility function might include variables not used in making choices in real markets. Economic theory has nothing to say about this matter; hence, whether choice processes are the same in real and hypothetical markets is an empirical issue, although there is now a growing body of evidence to suggest that choice processes can be very similar in both types of markets (e.g. Louviere and Swait 1996).

The ultimate purpose of estimating the choice model is generally to obtain unbiased estimates of the taste parameter vector β, which contains marginal utilities of attributes. While specification analyses and model fitting are the same whether the data are revealed preferences or stated choices from an experiment, it should be noted that choice experiments will often permit the identification of parameters (e.g. higher-order effects, interactions between variables, individual difference variables) that cannot be identified in RP data due to lack of variability, limited range or similar conditions. What is or is not identifiable in SC models will depend greatly (but not completely) on the experimental design, but it is generally the case that they will permit richer specifications than can be supported by RP models.

A component of the utility function that is commonly ignored in specification analysis is the error term. The most commonly applied choice model, the MNL, assumes that the error term is IID across alternatives and individuals. It is becoming more common to see more complex specifications of the covariance matrix of the error distribution, through models such as the Multinomial Probit (MNP) and Nested MNL. However, there is growing indication that the behavior of the error term in choice models may itself be explained in terms of covariates that impact the magnitudes of elements of the covariance matrix. For example, Swait and Adamowicz (1997) demonstrate that the complexity of the decision situation influences the magnitude of the variance of the error term and Cameron and Englin (1997) show that the variance of welfare measures depends systematically on the prior experience the respondent has with the good. We refer the reader to Hensher, Louviere and Swait (1997), who discuss this issue in greater depth.

As pointed out by Ben-Akiva and Lerman (1985) more than a decade ago, the specification of a choice model does not only involve the taste parameters β, but also the *choice set* from which the choice is made. For a number of reasons, including lack of recognition of this problem, almost all fields (econometrics, environmental work, marketing) that actively use choice modeling have only recently begun to examine the issue of choice set specification.[13] In transportation, seminal work was done by Swait (1984) and Swait and Ben-Akiva (1985, 1986, 1987a,b) in the modeling of choice set formation with revealed preference data; in Swait and Ben-Akiva (1985) it is shown that misspecification of the choice set yields biased taste parameter estimates. In the case of SC methods, the choice set formation problem is somewhat simplified since the overall choice set is under the control of the experimenter; nonetheless, individuals may eliminate particular alternatives that are offered from the choices that are considered. In the future, it would be advisable for environmental damage assessment measurements to account for choice set formation in both RP and SP data.

3.4 A Primer on Experimental Choice Analysis

It is our intention that this section provide an overview of the effort necessary to conduct an SC study. We provide references that permit deeper investigation of the method.

1. <u>Characterization of the decision problem:</u> As mentioned previously, this is a most important stage of the study. Through focus groups, literature search, interviews with experts, etc., the study team seeks to characterize the decision problem in terms that the decision maker understands. Specifically, we need to understand how individuals (i) become aware of the need to make the decision in question, (ii) define the dimensions of evaluation of the product or service, (iii) search for information on alternatives and attributes, (iv) construct choice sets, and (v) make decisions. These items are crucial in formulating a decision problem that is most akin to the decisions that individuals make in real life, when the selection problem of interest is one relatively familiar to decision makers. When the choice being studied is less familiar to the respondent, this stage maximizes our chances of communicating the desired information to him or her.

   We also seek to identify sources of individual heterogeneity (e.g. income, education, attitudes towards environmental issues) that could lead to important behavioral differences.[14]

   The basic outputs of this stage are four: (i) choice set size and composition, (ii) relevant attributes, (iii) individual differences and (iv) relevant sampling frame for the study.

[13] In the environmental valuation literature, for example, Peters et al (1995), Parsons and Kealy (1992) and Feather (1994) have explored this issue.

[14] Identifying and controlling for heterogeneity is important, but it is necessary to recognize that there will always be a chance that unidentified factors are influencing preference parameter estimates.

2. Attribute level selection: Based on study objectives and Step 1 information, the number and value of the levels for each attribute must be defined. This stage of the study is often conducted in parallel with Step 1, since even the language for communicating levels to individuals is often an issue. An important consideration at this stage is to be sure not to hamstring subsequent analyses due to an excessively limited range for the attributes.

   Commonly, attributes are identified from prior experience, secondary research and/or primary, exploratory research. Attribute identification procedures are discussed by Green and Srinivasan (1978, 1990), Louviere (1988a) and Timmermans and van der Heijden (1987). After identifying the attributes for a particular experiment, the analyst must assign values or *levels* to each attribute. These levels should be chosen to represent the relevant range of variation in the present or future market of interest.

   Though commonly presented in words and numbers, attribute levels may be communicated via pictures (static or dynamic), computer graphics, charts, etc. To the extent that visual (rather than text) representations of attribute levels are utilized, it is likely that respondents will perceive levels more homogeneously, likely leading to more precise parameter estimates in the modeling stage.[15] The tradeoff, of course, is that non-textual presentation of information is costly and (often) time-consuming to produce.

3. Experimental design development: Once attributes and associated levels have been determined, analysts typically use some form of orthogonal design to generate different combinations of attribute levels called "profiles" (e.g., Green 1974; Louviere 1988a). A profile is a single attribute level combination in a complete factorial combination of attribute levels (called a “treatment combination” in the statistical design literature). A “design” is a sample of profiles which have a particular set of statistical properties that determines the utility specification(s) that can be estimated (i.e. identified).

   Traditionally, linear model design theory has been used in developing SC designs (Louviere and Woodworth 1983, Louviere 1988b, 1994, Batsell and Louviere 1991, Bunch et al. 1993, Lazari and Anderson 1994, Kuhfeld et al. 1994, Huber and Zwerina 1996) despite the fact that most studies use nonlinear choice models to represent the data collected. (This issue is discussed extensively in the references above.) Such designs are available through published catalogs and specialized software.

[15] As part of a larger, on-going study of the external validity of SCM, Louviere, Anderson, Swait and Gray-Lee (1995) (LASG) compared SC models estimated from experiments in which (a) the attributes of alternatives were described purely verbally, and (b) attributes that could be visually described were so described, and the remaining attributes were described by identical verbal phrases used in experiment type (a). LASG found that the verbal and visual model produced identical utility estimates after rescaling to account for differences in variability; however, the scale of the utilities for the common verbally described attributes differed significantly from the scale of the utilities for the visually described attributes. The likely source of this difference is due to the fact that the subjects had to read and process the verbal information, form a mental image of the associated visual image and then integrate the resulting image with the remaining information. Subjects in the visual experiment could integrate the image directly and evaluate it with respect to the other attributes The verbal descriptions, therefore, induced more unreliability in the choices because of the need to read and form images, and insofar as different individuals perceive the images or interpret the words differently, variability in choice is to be expected. The visual condition eliminates both these sources of variability; hence, not surprisingly, choices in that experimental condition exhibited significantly less variability.

The reader should be aware, however, that the use of “canned” designs will often fall short of study needs. In many choice problems of interest there exist built-in constraints that must be satisfied to produce realistic choice scenarios. For example, a given level of one attribute may only appear if another attribute has a given value (e.g. a certain price may only be charged if the campground has a shower facility).

The overwhelming majority of experiments in which consumers express preferences for attribute level combinations presented one-at-a-time make use of orthogonal arrays called "main effects plans." A main effects plan is an orthogonal subset of the complete factorial which allows an analyst to estimate a strictly additive, "main effects only" (no interaction terms) utility specification. That is, one must know or assume that all interaction effects between attributes are not significant (Green 1974; Louviere 1988a); if this assumption is not satisfied, the use of main effects models can lead to unknown and potentially large bias in the utility parameters that are estimated. Main effects plans are unlikely to be appropriate in many choice contexts. Their popularity seems to arise from (i) a perceived need to model individuals to avoid aggregation biases, and (ii) insufficient training in statistical design theory, which results in overreliance on design catalogs and computerized design generators.

4. <u>Questionnaire development:</u> The questionnaire is usually a paper and pencil task that is either self-administered or presented through an interviewer.[16] While its main content will be one or more choice scenarios[17] through which the respondent will be guided, it may also include sections requesting sociodemographic, psychographic, attitudinal and past behavior data.

This last item (past behavior data) may be of particular interest if we intend to combine RP data with SC results. We not only have to collect information on what the individual actually did (e.g. where he/she fished), but also what other alternatives were considered, and if necessary, the characteristics of both chosen and non-chosen alternatives.

As in the contingent valuation literature, or indeed in any form or survey based research, pre-testing of the questionnaire is a necessary component of the research program. However, differently from CVM-type tasks, SC experiments require that the analyst define how many choice scenarios (i.e. replications) each respondent will be asked to do. While there are no hard and fast rules, the analyst must balance respondent learning and fatigue against efficient use of the respondent. There is contradictory evidence about

[16] However, a number of marketing and transportation studies have used video presentations, multimedia techniques in general to elicit stated choices. The vehicle for scenario presentation is more a function of available resources than any other consideration.

[17] Often SCMs create a form of panel data arising from the set of choice tasks presented to the respondent. For example, 8 or 16 choice sets may be presented to the individual. In contrast, traditional contingent valuation instruments present typically present a small number (one or two) choice sets. Double-bounded contingent valuation presents multiple choice sets, but the structure is somewhat different as the focus is on the bid or price variable. If a respondent selects "yes" to a willingness to pay question, he/she is then requested to answer a question involving a higher bid (Hanemann et al, 1991). This approach to constructing a panel is somewhat different than the approach taken in SCMs since in SCMs more than just the bid changes between choice sets. However, in both approaches the issue of how the early choices affect the later choices arises. In the double bounded contingent valuation literature Cameron and Quiggin (1992) and Alberini (1995) examine the issue of statistical relationships between responses. We later examine this issue in SCMs.

the impact of the task length on the quality of data provided by respondents (see, e.g., Brazell and Louviere 1997; Swait and Adamowicz 1997).

As a practical matter, we usually submit a respondent to about eight choice scenarios (see Carson et al. 1994a). However, we may give as little as one and as many as sixteen (or very occasionally, even thirty-two) scenarios to an individual. There has been little systematic analysis of the impact of varying the number of choice alternatives on individuals (though see Brazell and Louviere 1997; Swait and Adamowicz 1997). Furthermore, there is little analysis that we are aware of on non-response bias in SCMs, either item-non-response or survey-non-response. It may be the case that more complex, taxing survey designs will result in increased item non-response.

The actual means of dividing (or *blocking*) a design into manageable subsets of scenarios is easily accomplished. One can randomize the scenarios, then subdivide the reordered design to obtain subsets of desired size. Alternatively, one can generate a design that contains the blocking factor as an attribute with as many levels as there are blocks. If the blocking factor is orthogonal to all other design columns, the resulting blocks will have the desirable property that all levels of all attributes will be present in every block. However, this property may come at the price of a larger design than permitted by other considerations, in which case a tradeoff must be made.

5. Sample sizing and data collection: The usual considerations of desired accuracy levels versus data collection costs must guide definition of sample sizes. In addition, if we are estimating models that account for individual differences, we may have to impose minimum sample size requirements within segment to enable accurate predictions within segment. We refer the reader to Ben-Akiva and Lerman (1985), Daganzo (1980), and Cosslett (1981) for further discussions about sample sizes and sampling techniques for RP choice modeling. While the general principles discussed in these references are applicable, in SC experiments total sample size will be further affected by the total number of choice scenarios and the number of choice alternatives in a given scenario.[18]

6. Model estimation: The econometrics, transportation, marketing and resource economics literatures abound in examples of stated choice model estimations. The most common model estimated has been the Multinomial Logit (MNL), and the most common estimation criterion is maximum likelihood. However, there are also examples of other choice model specifications (e.g. Multinomial Probit, Nested MNL), as well as of other estimation criteria (parametric as well as non-parametric), being applied to SC data.

   If data fusion (i.e. combination of multiple data sources) is being performed, the estimation may involve both revealed and stated preference data. This may or may not require the development of specialized estimation software.

   It is less common to see, but completely viable to develop, models with other response variables. Olsen and Swait (1997), for example, collect both choices and quantities purchased from respondents in a SC task. They employ a Poisson Hurdle model to analyze

[18] Bunch and Batsell (1989) demonstrated that with as few as six respondents per choice scenario, the asymptotic properties for maximum likelihood-based inference are satisfied.

the quantity dependent variable. An example of joint analysis employing contingent valuation data is Cameron (1992).

7. Decision Support System (DSS) development: This step is quite idiosyncratic and specific to each study. However, in general, there is the need to embed the estimated choice model in a computerized tool that enables analyses to be easily performed. Though seemingly trivial, this step can be essential in making results (more readily) accessible to non-technical parties. This transparency may be of special interest in the case of environmental damage assessments, when increased access may translate into increased acceptance and confidence in the analysis results.

3.5 SCM Practice: Survey Design, Information Presentation, Context Effects, Correlated Observations and Capturing Temporal Behavior

Prior to examining a number of issues regarding SCM practice, we feel it appropriate to introduce the concept of the *scale* of the utility function (2). As we shall see, the scale (proportional to the inverse of the variance of the error term in the utility function) is a key quantity in understanding how categorical response models differ from the more familiar General Linear Model. It also affects interpretation of many of the issues to be subsequently discussed in this section.

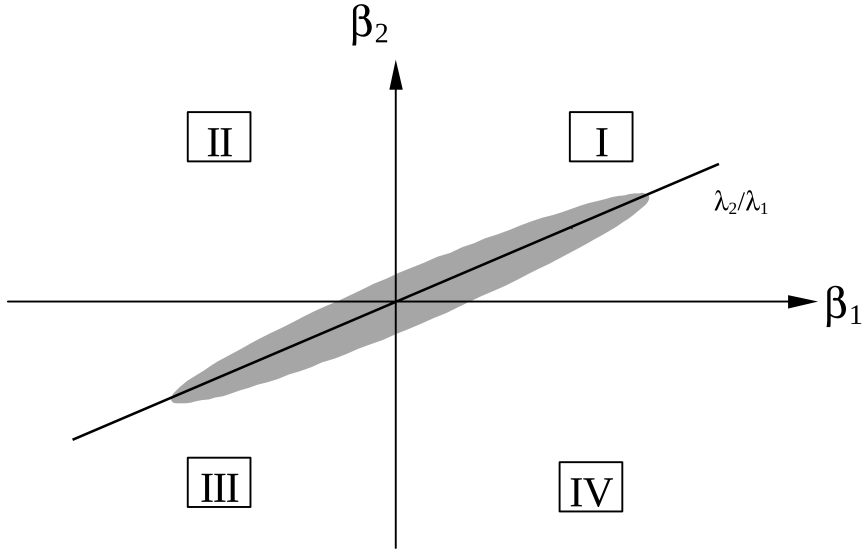
To derive operational models from utility function (2), assumptions are made concerning the joint distribution of the error terms. For expositional purposes only, assume that the errors are IID Gumbel (or Type I Extreme Value) with scale factor l>0, where $s^2 = p^2 / 6\, l^2$. Then, choice probabilities are given by the familiar MNL model:

$$P_i = \exp(\,lbX_i) \Big/ \sum_{j \in C} \exp(\,lbX_j) \qquad (6)$$

where we have assumed a linear-in-the-parameters specification for the deterministic component of the utility function.

Expression (6) illustrates that the scale factor and the other parameters of the MNL model are always in the multiplicative form ($l^{\cdot} k$), where k is some parameter vector, so it is not possible to identify the scale factor *within* a data source. Nonetheless, the scale factor affects the values of the estimated taste parameters: the larger (smaller) the scale, the bigger (smaller) the coefficients. Because of this confounding of scale and taste weights, *one should never directly compare the coefficients from different choice models* and conclude that one is larger than another. After all, what is causing the observed difference, scale or weight or both? In fact, note that even if two choice data sources are generated from the *same utility function* (hence have the same taste weights b), but have different scale factors l_1 and l_2, the estimated parameters will look different because in one case they'll be ($l_1 b$) and in the other ($l_2 b$). If both tastes and scales can vary between two data sets, parameter comparison is not a matter of eyeballing two lists.

Figure 3: Visual Test for Parameter Vector Equality Across Two Preference Data Sets



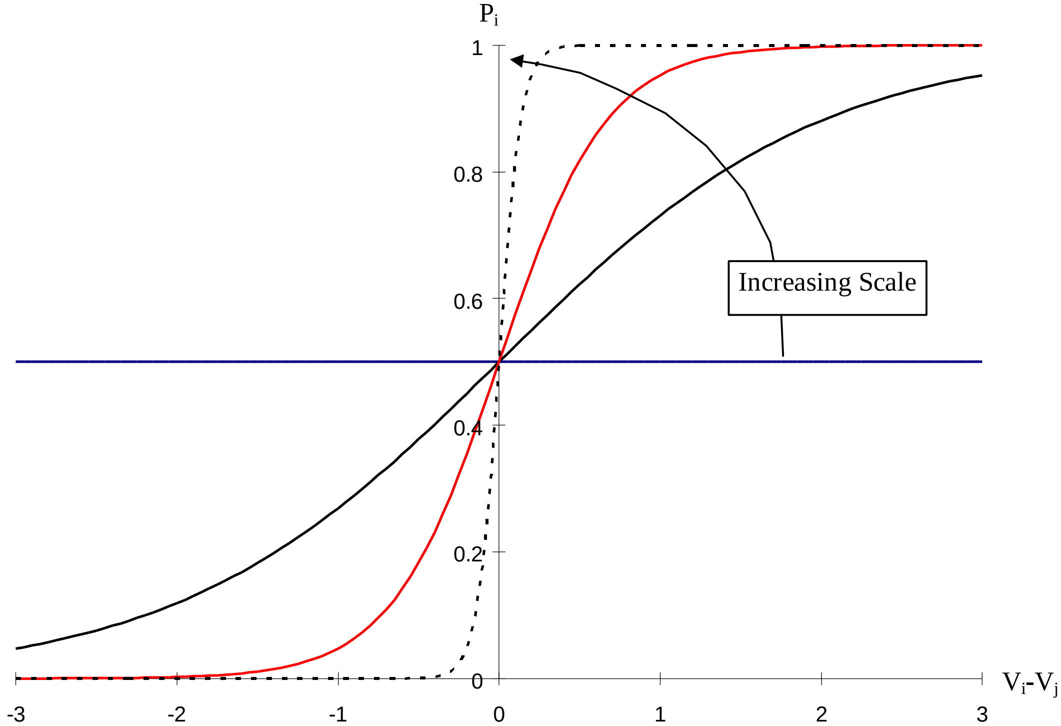
Return a moment to the comparison of two data sources that we believe to reflect the same tastes, but can have different scales. (Such as might be the case, for example, when combining RP and SP data.) The question of interest is whether $(l_1 b_1)=(l_2 b_2)$. Rearrange this expression as $b_1=(l_2/l_1)\,b_2$. If the two taste vectors are actually equal, this expression implies that a plot of one parameter vector against another should yield a cloud of points with (positive) slope equal to l_2/l_1, i.e. the relative variance of data set 2 to data set 1, going through the origin. Figure 3 shows how this might look, in general (these plots were suggested by Louviere and Swait 1996). To the extent that the cloud of points is too dispersed, or too many parameters have opposite signs in the data sources (implying points in quadrants II and IV), parameter equality between the data sets is unlikely.

Plotting parameter vector pairs in this manner is one of the easiest ways to looking for parameter equality in choice data. However, there are some problems with doing this: the main one is that the parameter estimates contain error, and the plots do not reflect this. Only a rigorous statistical test (Swait and Louviere 1993), can indicate whether parameter equality holds between data sets, after accounting for scale differences. Nonetheless, such plots are useful exploratory tools.[19]

But the influence of the scale factor of a data set is even more meaningful than it might seem at first. The key insight comes from the observation that the scale factor in the MNL model is *inversely* related to the variance of the error terms. This means that the higher the scale, the smaller the variance, which in turn implies that high-fit models have larger scales. As shown in Ben-Akiva and Lerman (1985) and repeated here in Figure 4, (1) when scale is zero the

[19] Another useful exploratory tool suggested by Louviere and Swait (1996) is to perform principal components analysis on multiple parameter vectors. If a single dimension results, this is supportive of parameter equality up to a multiplicative constant.

Figure 4: The Effect of The Scale Parameter on Choice Probability



choice probabilities are equal (in the binary case shown in the figure, both probabilities equal one half); (2) as scale grows from there, the choice model predicts more and more like a step function, which is to say, it becomes perfectly discriminating between the two alternatives in the graph. This effect generalizes for more than two alternatives.

The implication of these observations is that, differently from other model forms (and particularly the OLS form with which most analysts are acquainted), in choice models the taste weights and characteristics of the error terms are intimately (even inextricably!) related. Recent research has suggested that with choice models we must begin to think of the variance (equivalently, scale) as an integral part of the specification, not a nuisance parameter.

This relationship between mean and variance in the MNL model also occurs in other choice model forms: the NMNL and the MNP both exhibit this exact same effect.

Consider now the general issues of SC model reliability, as affected by survey design, information presentation, context and so forth. Any of these may affect both the mean and variance of utility estimates, either by introducing bias or affecting the scale of the utilities. Conceptually, introduction of bias is undesirable and to be avoided by the researcher. However, in this SCMs are no different from combinations of other elicitation methods and analysis techniques. Where the reliability of SCM is potentially different from other techniques is that certain practices which differentially affect scale should be accounted for. For example, Swait and Adamowicz (1997) show that choice task complexity and cumulative cognitive burden affect scale in a number of different data sets they examine.

As we see from Figure 4, higher scale factors correspond to steeper choice probability functions, which in turn means that there is improved discrimination between the preference for different alternatives. In the context of policy analysis, *ceteris paribus* these steeper curves are more preferable. Thus, the analyst must employ extant research, intuition, experience, etc., to produce survey instruments, present information and specify context in such a way as to increase scale. While this may seem vague advice to the reader, it reflects the current state-of-the-art.

The same type of understanding of the relationship between mean and scale in categorical response models also applies to the evaluation and testing of the external validity of SC experiments. In direct use applications of these methods, where an equivalent RP data source is available, parameter equality tests between data sources can be conducted while accounting for scale differences between sources (though these are not all in the environmental economics area, see, e.g., Ben-Akiva and Morikawa 1991; Swait and Louviere 1993; Swait, Louviere and Williams 1994; Adamowicz et al. 1994, 1995, 1997a,b; Hensher, Louviere and Swait 1997). The literature cited has reported a majority of instances in which SC experiments (designed using the general steps outlined in Section 3.4) are able to produce parameter estimates that are equal to RP data sources, up to scale.

3.5.1 SC Survey Design

As previously discussed, SCM surveys are the vehicle by which SC experiments are implemented. By "survey" we mean any vehicle by which scenarios and other questions are presented and/or asked, and responses and/or answers obtained. Thus, a survey can be a "paper and pencil" presentation of questions and choice sets, or it might be a full-blown multimedia event in which full-motion video and real-time audio and/or virtual reality are used to simulate different experimentally designed environmental scenarios as realistically as possible. Similarly, a "survey" can be administered by mail, personal interview or even via the Internet or other integrated network, like a dedicated cable channel.

Basically, SCM surveys typically consist of a glossary or similar section which provides basic information about the context, the attributes, the levels of the attributes, etc. The choice task itself is preceded by a set of standardized (for that task) instructions to subjects regarding the task, its objective, its context and how to respond to the scenarios. The task itself follows the instructions, and depends largely on the decision to be simulated and the research objectives. Thus, a task can be as simple as a yes/no response to scenarios containing single alternatives presented one-at-a-time, pairs of alternatives plus the option of not choosing, multiple alternatives with a no-choice option, and many variations thereof. In general, the task should be designed to simulate the actual choice and choice context as closely as possible. The scenarios typically appear in a random order, but there are no research results to provide guidance as to whether the ordering can be a) a single random ordering for all subjects, b) multiple random orders to which subjects are assigned, or c) separate random orders for each subject. Recently, however, Brazell and Louviere (1997) found no differences due to order, even in very lengthy tasks, except for differences in reliability: generally, shorter tasks were more reliable, but unreliability increased slowly with task length for numbers of scenarios that would be considered in applications (i.e. eight to 48). Order had no systematic effect.

3.5.2 Information Presentation

Information in SCM surveys refers to the alternatives and the attributes. This information can be purely verbal, purely visual, a combination of verbal and visual, multimedia, etc. The way in which information is presented depends on the problem and research objectives, as well as the resources available. The more realistically scenarios can be depicted, the better, which means that generally speaking, the trend has been increasingly to move in the direction of multimedia events when high levels of accuracy and reliability are demanded.

The way in which the information is presented also can vary substantially: verbal attribute information can be presented as a list of bulleted or highlighted items with the levels represented by short descriptors. Alternatively, such information can be presented in paragraph form. Generally, a picture or a graphic rendition is worth many words, especially if the words can be interpreted differently by different subjects. Thus, the trend has been to provide pictures, graphics, video, audio, etc.

3.5.3 Context Effects

A variety of context effects have been found in judgment and decision making experiments (Piattelli-Palmarini 1994). In fact, it is fair to say that there is a veritable academic cottage industry in finding and naming context effects. If one accepts the published results at face value, one would be forced to conclude that there was little hope of generality in any human decisions. In the case of SCM experiments or contingent valuation, however, there is ample evidence that models cross-validate well and produce utility estimates that are consistent with estimates obtained from parallel RP data. In the SCM literature this has been explored by Louviere and Swait 1996 while in the environmental valuation literature this issue has been examined by Carson et al. (1996).

That is not to say that there are not context effects in SCM experiments; rather, it is to say that there is little evidence that, if they do exist, they have much impact on model parameters or forecasts. Indeed, in a review of a number of studies in judgment and decision making in several disciplines, Oliphant et al. (1992) concluded that there was little evidence for context effects in choice experiments. More importantly, due to the fact that virtually no published research studies used double-blind techniques (in which the interviewers have no knowledge of the research objectives and/or hypotheses), it is unclear what can be said about results thus far obtained.

Framing is one of a variety of context effects that have been reported in the judgment and decision making literature (see Piattelli-Palmarini 1994). It refers to the way in which the choice or decision context is worded, the situation invoked and/or the reference used (e.g., is it a loss or a gain, is it positive or negative, etc.). Generally, speaking, research has shown that framing affects the choices individuals make and the parameters of the functions which drive same. What is less clear is whether framing induces real process differences or merely influences the variance of the random component of utility. Recent results (e.g. Olsen et al. 1995; Louviere and Swait 1996; Louviere, Fox and Moore 1993) suggest that many context effects may in fact only impact variance and not utility parameters. Nonetheless, it should be noted that impacting error term variance does impact outcome predictions (as the scale of the error term increases, choice probabilities also increase). Thus, the issue is far from settled, and is the subject of ongoing research. What can be said is that there is evidence that framing has an effect on the output of choice experiments, but it is unclear how serious such effects might be in estimating welfare measures.

Until such research is conducted, however, it seems to us that framing conditions imposed in tasks that will be used in environmental damage assessments situations must be evaluated principally with respect to potential difficulties introduced to policy analysis. For example, if individuals' tastes are impacted by whether a beach closure is due to a natural occurrence (e.g. "red tide") or to a shipping accident (e.g. an oil spill), it may be extremely important to evaluate the impact of the framing of the event on the ensuing welfare measures.

3.5.4 Correlated Responses

As mentioned above, SC tasks almost always expose respondents to multiple choice scenarios. Traditionally, parameter estimation has proceeded under the assumption that the multiple responses of each individual are independent, mainly because of significant difficulties associated with relaxing this restriction. The biggest source of difficulty with handling repeated measures arises from the need to develop a structural model of the relationship between sequential responses from a single subject. In particular, a specific model of learning and fatigue must be adopted, which must incorporate effects on both means and variances due to past responses. Clearly, several alternative models of both processes are possible; it may be difficult to achieve consensus on the general desirability of one formulation over another. This type of problem has parallels with the dynamic choice model literature (e.g. Heckman 1981), where habit formation, inertia and learning must be considered.

Research by Brazell and Louviere (1996) and Swait and Adamowicz (1997) supports the hypothesis that rather than correlation, the problem with multiple responses may instead be non-stationary variance. The latter, for example, demonstrate that the variance in SC data collected in the format described above is a function of decision complexity and cumulative effort expended (i.e. learning and fatigue effects can occur, but are assumed to affect the variances of sequential responses).

While improved precision may be of interest generally, in this case it will entail justifying the use of one particular correlation structure vis-à-vis *many* other plausible possibilities. This may be one of those cases in which the remedy may be more controversial than the original problem. We instead urge utilization of alternative methods to recalculate standard errors for possible misspecification effects:

1. ***Simple Correction:*** multiply the calculated errors output by a maximum likelihood estimation program by the square root of the number of replications seen by each respondent (this assumes all respondents see the same number of scenarios), then recalculate the asymptotic t-ratios. The rationale behind this method is that the degrees of freedom are being overestimated because we do not have N·R independent observations, where N is the number of subjects and R is the number of repeated measures from each subject. *Conservatively*, we can assume that we have only N independent observations. This leads to a multiplicative correction factor of $R^{1/2}$ to be applied to the standard errors output by an estimation that assumed all N·R observations to be independent. To our knowledge, this was first suggested by Louviere and Woodworth (1983). This method is useful when employing estimation programs that do not implement the next method;

2. ***White's Quasi-Maximum Likelihood Method:*** If the data generation process is misspecified and we nonetheless use ML to estimate process parameters, given a number of standard assumptions, White (1982) showed that the resulting estimates are asymptotically consistent and normally distributed, with covariance matrix given by
$$\Sigma_Q = \mathrm{F}^{-1}\Gamma\mathrm{F}^{-1}, \tag{7}$$
where F is the Fisher information matrix and Γ is the cross-product matrix of first partial derivatives of the log likelihood function. At the ML estimate, if the dgp is correctly specified, Γ=F and (7) reduces to the inverse of F. Hence, the quasi-maximum likelihood (QML) estimator reduces to the ML estimator when the dgp is correct. If the dgp is somehow misspecified (e.g. it does not consider repeated measures within subject), expression (7) is the appropriate means of calculating standard errors. These will be larger than those yielded from assuming independence among repeated measures within subject, just as in the first method. (Note that White 1982 also provides a misspecification test based on the difference between Σ_Q and F^{-1}.)

This second method, while more difficult to implement than the first, has the advantage of being theoretically rigorous and well-accepted in the econometrics community. To our knowledge, use of the second method has not been reported in the SP literature.

3.5.5 Temporal Concerns

SCM has not been applied to study temporal changes in behavior or sequential decision events. To do so would require a panel approach in which appropriate samples of subjects were repeatedly exposed to different experimental treatments, and their choices recorded. In theory this could be done, especially given recent advances in integrated networking and multimedia technology; however, at present it would be quite expensive. Generally speaking, however, it is at present difficult to imagine how one would structure a panel study that would be sufficiently rich to capture the dynamics of real decision environments. Thus, for the foreseeable future it is likely that SCM will be limited to one-off or discretely staged applications.

It is our opinion that temporal and dynamic effects are presently best handled by panel monitoring using RP data although methods for repeated stated preference surveys may be possible if resources for implementation are available.

3.6 Advanced Issues In Stated Choice Methods and Random Utility Theory

3.6.1 Fixed Alternatives

Modeling Participation Decisions and Other Fixed Alternatives

It is not uncommon for an event of interest (e.g. an oil spill) to not only affect decisions involving *which* alternative to choose (e.g. which beach to go to), but also affect whether or not to *participate* in some form of activity (i.e. whether to go to a beach at all). It is straightforward to design SC tasks to include the participation decision, and model the data utilizing a specification such as the Nested MNL. Figure 5, for example, depicts a hypothetical choice scenario for camp site selection with three designed campsites and three fixed alternatives (so-called because their attributes are not varied according to the statistical design). Of the fixed alternatives, two represent specific aggregate alternatives ("campsite off Prince Edward Island," and "hotel, motel, lodge or cabin") and the third captures the participation decision.

The important issue in handling the participation decision with an SC task is to make that alternative clearly understood in the same way to all respondents. This objective will usually be accomplished by making the fixed alternative description as specific as possible. This comment is equally valid for other fixed alternatives.

The Choice Not to Choose

The likely number and proportion of individuals who choose from offerings in a particular category of goods or services at any point in time depends on their attractiveness, expense and risk, as well as stage in the product life cycle (or degree of decision maker familiarity with the category). Expensive, risky new alternatives or technologies frequently exhibit high levels of non-choice in the introductory and early expansion phases of their life cycles. Thus, it is important to recognize that in such cases one wants to model not only the probabilities that consumers will choose something, but also the probability that they will choose nothing.[20] More generally, however, one should design SC experiments to allow one to observe and model non-choice because it's such an obvious element of real market behavior.[21]

[20] Clearly, if one wants to model only choosers, one should exclude individuals with zero purchase probabilities.
[21] In contrast, no theoretically acceptable way has yet been proposed to observe and model non-choice in traditional conjoint analysis experiments.

Figure 5: Example Choice Scenario with Fixed Alternatives

FEATURES:	Campsite 1	Campsite 2	Campsite 3	Campsite Off Prince Edward Island	Hotel, Motel, Lodge or Cabin	Not Go Camping and Stay at Home
... ...	... ...	... ...	... ...			

⇩ ⇩ ⇩ ⇩ ⇩ ⇩

Suppose only the campsites shown above were available to you. Which alternative would you choose?

Check one and ONLY ONE box:	☐	☐	☐	☐	☐	☐

Fortunately, SCM provide a theoretically acceptable way to measure and model non-choice based on the theory of consumer demand, and in particular in the random utility models of discrete choice. That is, probabilistic discrete choice theory recognizes that many elemental goods are indivisible (McFadden 1974); such goods are chosen in discrete units. Louviere and Woodworth (1983) provide the conceptual and methodological foundation for designing and implementing SC experiments for discrete goods, in which they treat non-choice as simply another discrete option faced by consumers (see also Anderson et al. 1992).

3.6.2 Treating Heterogeneity in SC Experiments

Treatment of taste heterogeneity in SCM is identical to that developed for RP methods. For example, in order of increasing generality, one may

1. define *a priori* segments (e.g. income, location, experience, frequency of use) and interact them with design attributes to capture differential attribute sensitivities (see, e.g., Louviere 1988a,b, Batsell and Louviere 1992, Kaciak and Louviere 1990);

2. estimate a latent class model, which is a special case of a random parameters specification, in which a discrete number of support points are hypothesized (see, e.g., Swait 1994);

3. estimate a random parameters model, which postulates continuous distributions for parameters (see, e.g., Layton, 1996).

Random parameters models seem more appropriate for better characterization of mean utility levels through improved treatment of population heterogeneity. From the point of view of policy and welfare analysis, however, the degree of functional flexibility afforded by random parameters model may be offset by the reduced interpretability of the causes of heterogeneity, hence lesser understanding of the distributional impacts of alternative policy scenarios. Since there may be interest in knowing how different socioeconomic groups may be affected by policy options, it may be that the simpler *a priori* segmentation approach may be most applicable. Swait (1994), however, developed a latent class model in which class membership can be a function of sociodemographic and psychographic characteristics. Hence, the latent class approach may represent a workable compromise between capturing behavioral differences between subjects and aiding interpretability of impacts on *a priori* groups.

Historically, in the marketing literature, treatment of heterogeneity in (ratings) conjoint models began by developing individual level conjoint models,[22] which in a way is the ultimate means of avoiding aggregation bias and identifying potential segments. However, Dawes and Corrigan (1974), Wainer (1976), Anderson and Shanteau (1977) and others demonstrate that utilities estimated from additive, individual-level models can be very unreliable and biased. We do not generally recommend the employment of this method.

3.6.3 Violations of the Independence of Irrelevant Alternatives Property

Louviere and Woodworth (1983), and Batsell and Louviere (1991) discuss how one can design SC experiments to test violations of Independent and Identically Distributed (IID) error terms. (Hausman and McFadden (1984) and McFadden (1987) provide actual test methods.) If violations are found, the same methods available for RP data can be applied to SC experiments. This type of test is of special interest in the case of the MNL model, known to be characterized by the Independence of Irrelevant Alternatives (IIA) property, which is a direct reflection of the IID nature of the error terms in the utility function. Despite the recent availability of estimation software for more advanced model forms, the MNL is still the most widely used formulation for discrete data, so the issue of testing for IIA violations is still of interest.

If IID violations are found, the researcher can utilize more flexible specifications such as GEV-based and Multinomial Probit discrete choice models. In the case of GEV models, such as the Nested MNL, patterns of unobserved similarities between alternatives are described in the form of trees which assume that within clusters of alternatives the IID structure holds, but that between clusters there exists a correlation pattern. Thus, the analyst's modeling task is augmented by the need to specify the correct tree, which leads to a certain amount of exploratory work. MNP models permit more general covariance structures than GEV models.

[22] Standard practice in conjoint analysis is to collect sufficient responses from each subject to permit calibration of individual-specific models, hence the term "individual-level models." Since each model is based on a few observations (usually, as few as one can get away with) and between subject regularities are not exploited, parameter estimates tend to be poorly estimated. Nonetheless, it is common practice among conjoint analysts to cluster respondents on the basis of similar parameters to find segments of interest (price-sensitive segments, brand-loyal segments, etc.).

Bunch (1991), however, has shown that because the MNP model is identified by differences in utilities, not all covariance structures are identifiable. In fact, he mentions several instances in the econometrics and transportation literatures in which non-identified covariance matrices were estimated for MNP models. Bunch's work shows that MNP covariance matrices corresponding to tree structures (as in GEV models) are identifiable, hence the claimed superiority of MNP may be more perceptual than factual.[23] Hence, given the greater availability of NMNL estimation software and that model's greater ease of estimation vis-à-vis MNP models, we recommend that NMNL models be employed whenever IID violations are detected.[24]

As far as designing experiments to permit detection of IID violations, Louviere (1988a,b) proposed that a sufficient condition to detect IID violations is to design SC experiments to insure that all attributes of all choice alternatives are orthogonal to one another. This design approach satisfies the conditions needed to test IID violations because it is consistent with a very general stochastic choice model developed by McFadden (1975) called Mother Logit.

Designing SC experiments to be consistent with the Mother Logit model not only allows one to test IID violations, but also to include cross-alternative effects in the model to take the violations into account. That is, Mother Logit consists of a system of simultaneous equations for all choice alternatives in which the attributes of all alternatives are free to enter the utility functions of any alternatives.[25] If the IID error assumption is satisfied, the only attribute effects that will be significant in any utility functions are "own-attribute" effects. That is, only the attributes of choice alternative A can affect the utility of alternative A; the attributes of alternative B cannot influence choices of A. For example, if preference heterogeneity is the primary source of IID violations, a well-specified model with appropriate individual difference effects should eliminate the violations and provide insights useful for segmentation and targeting. For example, Swait et al. (1993) found highly significant IID violations, but once they incorporated several individual difference measures in their models, no IID violations were significant in any of the product categories they studied.

3.6.4 Combining Multiple Choice Data Sources

As previously discussed, a major advantage of SC experiments is that models estimated from the choice data they generate are consistent with random utility theory. This matters because traditional conjoint analysis responses (e.g. ratings, rankings) and (associated) models are generally inconsistent with random utility theory, as usually analyzed; more importantly, they do not correspond to actual behavior in the sense that ratings, rankings, or the like have no obvious parallel in real markets. On the other hand, choices made by consumers in SC

23 "In fact, users of probit must in the final analysis make modeling assumptions which are analogous to choosing among various alternative tree structures in the nested multinomial logit or tree extreme value models." (Bunch, 1991, 11)

24 Of course, nothing keeps the researcher from starting out with the more complex model form in the first place, and reducing to the MNL model if warranted by the data.

25 McFadden's (1987) IIA specification test for the MNL model suggests that the theoretically correct way to include cross-effects is as the square root of the variable included in the "own" utility function.

experiments reveal preferences in a manner analogous to the way choices in real markets reveal preferences. Hence, choices in real markets and SC experiments can be compared because both are the outcome of processes consistent with random utility theory.

Thus, Ben-Akiva and Morikawa (1991), Swait and Louviere (1993), Adamowicz, Louviere and Williams (1994), Swait, Louviere and Williams (1994) and Louviere, Fox and Moore (1992) jointly and simultaneously estimate choice models from SC experiments and marketplace choices by taking account of differences in the variances of the random utility components in both data sources. That is, the unit of measurement of utility scales in discrete choice models is inversely proportional to the error variability in the choice data, and it is unlikely that market and SC choices have the same variability. Yet, if both sources of data are the outcome of a common choice process which differs only in the variability of the random component, utility parameters estimated from each data source should be proportional. A Full Information Maximum Likelihood estimator for the relative scale parameter can be developed (see, e.g., Swait and Adamowicz 1997).

However, a simple method that yields consistent, but not efficient, parameter estimates was formulated by Swait and Louviere (1993), who showed how to estimate the ratio of the scale units in two or more data sets and test for a common choice process. The Swait and Louviere approach to estimate the appropriate scale ratio to transform SC parameters to scale with marketplace choices is as follows:[26] a) Measure and code attributes common to both data sources in the same way. Code effects which are unique to one or the other data source (like alternative-specific intercepts) in the appropriate way, zeroing out such columns in the other data source. b) Rescale one data source by multiplying the design matrix columns by an initial constant within the range expected to contain scale unit ratio. c) Combine both data sets into one data set. d) Estimate the overall model and obtain its associated log likelihood value. e) Repeat (b) to (d) to conduct a grid search of the region in which one expects to find the true scale ratio. The log likelihood function has a unique maximum at that value of the scalar multiplier (which will be the relative scale of the two error terms) which best estimates the scale ratio. The resulting model utilities for the joint model will be scaled relative to the units of the reference data set (e.g. SC rescaled to market as reference).[27] While this procedure is convenient for those who have access only to MNL estimation software it should be noted that this method will generate underestimates of the standard errors.

3.7 Stated Choice Methods and Environmental Choice Analysis

3.7.1 Applications

Stated choice methods have been widely employed in marketing and transportation to analyze consumer choices of products, modes of travel and a variety of other items. Hensher (1994) provides an overview of SCM as they have been applied in transportation; Louviere (1994)

[26] This method is specific to the case of pooling two data sources, say SP choices and RP choices. It becomes impractical with three or more data sources.
[27] Parameter equivalence after accounting for scale differences can be tested using a likelihood ratio test.

does the same for applications in marketing.

While the use of attribute based stated choice techniques[28] (or the sub-set of stated preference techniques that consist of alternatives defined by attributes) in environmental valuation is relatively recent, there have been a number of noteworthy examples. Rae (1983) employed SCM type techniques in the analysis of benefits from air quality improvements. Lareau and Rae (1988) studied the value of odor reductions using a type of SCM model. Mackenzie (1993) employed SCM type techniques to examine tradeoffs between attributes of recreational hunting experiences. Opaluch et al. (1993) used paired comparisons in an SCM framework to analyze hazardous waste siting decisions. Viscusi et al. (1991) employed SCM type techniques in analyzing health risk tradeoffs. Goodman (1989) examines housing attributes in a stated preference conjoint framework (see also Freeman, 1991).

Adamowicz et al. (1994) employed a SCM approach to value the impact of a water resource development. This model was constructed to examine recreational site choice. They also examined a revealed preference model of site choice, and combined the two approaches. Among the interesting findings from this study was the fact that the revealed preference and stated preference models were not significantly different (once differences in scale factors were accounted for; see Section 2.6.4). Boxall et al. (1996) compared SCM techniques with contingent valuation in the case of the improvement of a recreational hunting site. Adamowicz et al. (1997a) examine joint stated and revealed preference data in an environmental valuation context as well. In this case they consider the impact of using perceptions versus objective measures of attribute data in these joint models. The paper provides a more generalized view of the types of data available to study individual choices and the relationships between these different data structures.

While the examples above have focused on recreational site choice (or use values), Adamowicz et al. (1998) use SCMs to assess a habitat protection program or a passive use valuation exercise. They compare the stated choice results with a referendum contingent valuation experiment and they combine the information from both data sources in a joint estimation. The results suggest that stated choice methods are particularly useful in passive use cases when attributes are a focus of the analysis. Finally, Adamowicz et al. (1997b) provide an overview of SCMs as they are applied to environmental valuation.

3.7.2 Selected Issues in SCM Applied to Environmental Problems

Contingent Valuation versus SCM

Recent interest in SCMs has arisen in part in response to criticism of the Contingent Valuation Method (CVM) and in particular to the use of CVM in passive use value contexts (e.g. Green,

[28] Of course contingent valuation, a stated preference method, has been applied to environmental valuation for at least 3 decades.

1995; Desvousges et al, 1995).[29] It may be possible for some of the criticisms of CVM to be addressed by using SCMs. For example, SCMs involve a broader attribute based perspective, where the attributes of the choice context in general are considered, rather than the specific details of the case at hand. Furthermore, the SCM approach focuses on tradeoffs over several attributes, rather than only or primarily on price (or payment). The repeated choice context of SCM allows for preference construction to occur and may be used to remove the influence of heuristics employed by respondents and reveal information provided in an actual tradeoff context (Swait and Adamowicz, 1997; McFadden, 1996). Furthermore, the repeated question nature of SCMs allows for internal validity tests and provides a response surface which may yield important information about the consistency of individual responses (Green, 1995). In the context of use values, significant evidence already exists on the validity of SCMs in predicting actual choice responses (Louviere and Swait, 1996; Louviere, 1996). This evidence, at least in part, addresses the concern that SCMs employ “hypothetical” questions and thus cannot be trusted to predict real world choices.[30]

However, there are still some issues that have arisen in the criticism of CVM that may be applicable to SCMs. Just as in CVM, careful experimental design, including the selection of attribute levels, choice contexts, survey design and implementation, and sampling methods are required for SCM tasks to provide the best results. In cases where respondents are making choices in unfamiliar contexts, SCM may outperform CVM if effort is made to describe the choice in a fashion that can, over repeated tasks, become a familiar tradeoff context somewhat similar to a market choice. An issue as yet unresolved in SCM is the degree to which respondent learning (preference construction) and/or respondent fatigue can be captured and assessed within the SCM task. Recent evidence suggests that learning and fatigue effects occur in choice tasks, however, the degree to which such effects are present is context dependent (Swait and Adamowicz, 1997). However, it is important to develop methods by which such effects can be detected and assessed within the analysis. Finally, as both Green (1995) and Desvousges et al (1995) point out, in the case of passive use values there are no obvious external validity tests that can be performed except perhaps to test SCM results with results from actual referenda. However, SCMs seem to offer more room for internal validity checks through the repeated choice context and the development of a broader response surface.

SCMs currently suffer from the fact that they have not been tested in an NRDA context and have not been put through the rigorous evaluation of this system. However, this should not be viewed as a criticism of the method; it is simply a consequence of the path of events. Indeed, several major critics of CVM have described SCM as a promising alternative (Desvousges et al, 1995; Green, 1992; McFadden, 1996).

[29] As discussed in the introduction to this document, SCMs can be considered a generalization of referendum contingent valuation.

[30] Evidence that SCMs can generate preference maps that are consistent with results derived from revealed preference data does not preclude the need for careful design (experimental design and survey design) in such exercises.

Measuring Gains versus Losses

An issue that has generated substantial controversy in the non-market valuation literature is the difference between valuations of gains and losses. Whether one adopts the notion of irreversible indifference curves (Knetsch, 1991), or the notion that large divergences between willingness to pay (WTP) and willingness to accept (WTA) compensation can be accounted for by differences in the degree of substitution possibilities available for the good or service in question (Hanemann, 1991), the fact remains that differences between these two measures can be substantial. Furthermore, when using contingent valuation methods in environmental contexts, measuring willingness to accept compensation is often very difficult and replacing such measures with the more easily derived willingness to pay measure results in undervaluation of the losses.

While there has been little formal analysis of this issue within an SCM framework, there are some interesting attempts to address the measurement of WTA. For example, Peterson et al. (1995) employ a pairwise stated choice method to elicit WTA and find that the method works remarkably well. Adamowicz et al. (1998) design their SCM experiment to include situations in which WTA is being elicited as well as cases in which WTP is obtained. This is one area in which the difference between traditional contingent valuation methods, which focus on a specific case in great detail, and SCM, which examine the universe of possibilities surrounding a class of goods or services, may be quite important. While individuals may have great difficulty accepting compensation (as presented in a contingent valuation experiment) for damages to specific ecosystems , they may have little difficulty in choosing one situation out of several options where their property taxes would be lower but the level of environmental quality would also be somewhat lower. Nevertheless, this is a challenging issue and little has been done on this topic within the context of SCM. Future research will reveal whether such choice methods are conducive to the elicitation of WTA measures.

Choice of Non-Market Goods and Services

As presented above, Stated Choice Methods have been used for valuing non-market goods and services, although these have primarily been measures of use value. Use values like recreational site choice are quite similar in nature to market products since the consumer faces market prices (access costs) and can choose from a range of alternatives, each with certain attributes or characteristics. Clearly, since SCM have been largely employed in the fields of marketing and transportation, the link between these methods and measuring use values in environmental contexts is strong. Methods like random utility theory have rapidly spread as the preferred method for modeling observed behavior on site choice (or travel cost models). Stated choice methods provide a method to "improve upon" these revealed preference methods and maintain the utility theoretic nature of the method.

While the linkage between SCM and use value measurement techniques is relatively well established (e.g. Adamowicz et al. 1994), there has been limited exploration of their use in measuring passive use values. The accepted method for measuring passive use values is a version of referendum contingent valuation (Mitchell and Carson, 1989). SCM could also be constructed in a binary choice, referendum fashion, thus preserving the anticipated incentive compatible nature of this method.[31] A recent paper by McDaniels (1996) illustrates the use of a simple stated choice method as a three alternative referendum that was actually carried out in Victoria, Canada. Similarly, the Adamowicz et al. (1998) paper constructs the choice scenario as a set of options for an individual to consider as if they were voting on selected "futures." While there has been relatively little actual experience with SCM in passive use value contexts, they appear to be promising as methods of analysis. However, further research on the design format, and their incentive properties, is required.[32]

3.8 Discussion

Our primary goal in this report has been to provide a behavioral basis for Stated Choice Methods (SCM) by emphasizing the random utility framework for models and measurement procedures. In random utility theory consumers try to maximize utility functions that contain random components, which implies that choice must be stochastic when viewed from a researcher's standpoint. More importantly, this view allows us to place SCM within mainstream microeconomic theory, linking SCM models to a larger family of probabilistic discrete choice models in econometrics. The latter link matters because no market behavior corresponds to that observed in traditional SP rating and ranking tasks. However, the links to real market behavior are not only clear in SCM, but one can combine and compare models based on SCM and real market choices.

We have also presented a simple outline of the steps of a SC study, and discussed a number of issues that arise in the application of SCM to the environmental valuation area. Many challenges remain to the widespread application of SCM to this area, but we feel that overcoming some of these has more to do with establishing adequate "comfort levels" in the relevant disciplines than with actual problems with SCM. Nonetheless, a good many issues, and certainly the more difficult ones, remain yet to be researched to our satisfaction.

4. Conclusion

In the course of developing this paper, it has become clear to us that the implications of switching to a resource based compensation method are more far-reaching than we first thought. Expanding restoration options to include non-monetary compensation implies that attribute-based methods must be used to develop measurement and evaluation tools. While

---

[31] There is a significant literature on the incentive compatibility of CVM (e.g. Mitchell and Carson, 1989; Carson et al. 1997). The most commonly cited finding from this literature is that the two alternative referendum approach to CVM is incentive compatible. The incentive compatibility of the this approach, however, relies on the respondent "believing" the scenario and the implications of the alternatives (Carson et al, 1997 and McFadden, 1994). Design issues are as important in this format as they are within other valuation formats.

[32] The degree of strategic behavior expected from multi-alternative referendum type questions, and the degree to which the influence of such strategic behavior can be "removed" in the analysis, is an unresolved question (Carson et al, 1997). However, it appears that the referendum format offers the best potential for incentive compatibility.

revealed preference (RP) and traditional CVM methods can sometimes be used to fulfill this requirement, they are not always able to do so. This is particularly the case with passive use values: RP and traditional CVM measures will either not be possible, or may not work very well (see Adamowicz and Swait, 1997).

Stated Choice methods (SCM) have been extensively used in product and service design problems in marketing, transportation and geography, among others. There is already developing a significant literature, contained in our reference list, showing the feasibility and benefits of applying them to environmental valuation. However, to our knowledge, with the exception of Adamowicz et al. (1998), Opaluch et al. (1993) and McDaniels (1996), SC applications have been limited to mimicking and supporting RP methods. In this report we have pointed out that resource based compensation valuation requires a level of flexibility in the data collection method that can most likely be better provided by SC than RP and CVM methods. Some extensions to existing SCM should be investigated, however, to decide how best to adapt, change and extend them to improve these methods for application in resource based compensation projects.

A further concern arising from the issues described above is that SCMs have been applied primarily to the examination of private goods while most, if not all, cases of environmental valuation and proposed resource compensation involve public goods. Due to the lack of research in this area, there are a large number of issues that require attention. Many of the concerns raised by the NOAA panel regarding the application of contingent valuation are applicable to SCMs. Issues of incentive compatibility, information provision, task complexity, sampling frame and other concerns raised by the panel apply to SCMs since SCMs are still survey based instruments attempting to elicit choices as if they had been made in "real" contexts (markets or referenda).

A research issue that has been raised in this review is the need for investigation of the impacts of survey complexity and response burden. While there have been some attempts to examine this issue in the SCM literature there does not appear to be any systematic attempt to assess cognitive burden as it relates to the number of attributes, the number of alternatives, the number of replications and a variety of other SCM design issues. SCMs, or attribute based choice tasks, involve a different type of complexity than CVM tasks as they tend to focus on categories of attributes identified as important for decision making. Also, repeated choice sets create a form of cognitive burden not typically found in CVM. However, CVM relies on detailed descriptions of the situation with less variation at the choice level. The cognitive requirements are undoubtedly different, but it is not clear that one is more burdensome than the other. Both methods strive for reality, and attempt to be able to predict responses that would have been made in a real market or choice situations. However, the methods employ different approaches to eliciting this information and predicting these responses.

Finally, there are several significant challenges associated with a movement to resource compensation, including the increase in complexity of the compensation determination problem. A further challenge will be to link rapidly developing economic methods with ecological service flow models in order to understand the indirect linkages between ecological services and direct use values. While the determination of resource compensation is a formidable task, the potential payoffs are high and the recent developments of Stated Choice Methods provide the tools necessary to take up the challenge.

Acknowledgments

The authors would like to thank David Chapman, Carol Jones and Curtis Carlson of the NOAA Damage Assessment Center, Resource Valuation Branch, for guiding us throughout the development of this document. We would also like to express our thanks to Richard Bishop, Trudy Ann Cameron, James Heckman, Alan Randall, Paul Ruud, Raymond Kopp and Stanley Presser for their review comments and suggestions.

Glossary

Conjoint (Choice): A method of preference elicitation that presents to respondents one or more sets of two or more alternatives (generally experimentally designed) and asks that they indicate their most preferred alternative. Such data are analyzed assuming that Random Utility theory is the underlying data generation process. Contrast with Conjoint (Ratings).

Conjoint (Rankings): A method of preference elicitation that asks respondents to rank several alternatives. This ranking can be examined using random utility models (e.g. exploded logit models) and can be construed to be consistent with random utility based revealed preference models. However, in actual markets consumers are not commonly observed ranking alternatives.

Conjoint (Ratings): Ratings-based conjoint is a preference elicitation method that has different forms. The most common is to present a respondent with a single product description (called a profile) and elicit their likelihood of purchase on a rating scale (say, 1=highly unlikely to purchase, 7=highly likely to purchase). Sufficient product descriptions are presented to each respondent so that individual level models are usually calibrated. The most common method of analysis has traditionally been OLS, which is inappropriate given its implicit assumption of interval scaling throughout, which is unlikely to hold in practice. Choice behavior is *not* elicited, so that subsequent market predictions must be done by assuming some relationship between the rating function and choice; usually, the highest rating product in the prediction set is assumed to be chosen. Ancillary analyses often utilize the individual level parameters as the basis for market segmentation schemes. Contrast with Conjoint (Choice).

Conjoint Analysis: See Conjoint (Ratings).

Factorial Design: All possible combinations of attributes and levels in an experimental design including main (linear) effects and interaction (cross) effects.

Factors: A set of attributes each of which may have 2 or more levels.

Fractional Factorial Plan: A fraction (subset) of the factorial plan, usually selected after making assumptions on the degree of interaction terms estimable in the model (e.g. a main effects only design).

Orthogonal Design: An experimental design in which attribute levels across alternatives are uncorrelated thereby providing unconfounded measures of the partworth utilities or attribute parameters.

Partworth utility: This is the total utility associated with a given level of an attribute. If $\boldsymbol{b}_k$ is the marginal utility of a unit of continuous attribute X_k, then ($\boldsymbol{b}_k X_k$) is the partworth.

Psychographic: Measurements pertaining to psychological attitudes, emotions, perceptions and so forth.

Stimuli: In the context of Stated Choice Methods, these are the attribute levels (price, distance, color, …) and context variables (whether the hotel/motel alternative is available in a campsite model) that constitute the pieces of information assumed to impact the decision maker's response and being manipulated by the experiment.

References

Adamowicz, W.L., T. Beckley, D. Hatton-Macdonald, L. Just, M. Luckert, E. Murray, and W. Phillips. 1997c. In Search of Forest Resource Values of Aboriginal Peoples: The Applicability of Non-Market Valuation Techniques. Forthcoming, S*ociety and Natural Resources.*

Adamowicz, W., P. Boxall, J. Louviere, J. Swait and M. Williams. 1997b. Stated Preference Methods for Valuing Environmental Amenities. in *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the U.S., E.C. and Developing Countries*. I. Bateman and K. Willis, editors. Oxford University Press, forthcoming.

Adamowicz, W.L., P. Boxall, M. Williams and J. Louviere. 1998. Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation. Forthcoming, *American Journal of Agricultural Economics.*

Adamowicz, W.L., J. Louviere and M. Williams. 1994. Combining Stated and Revealed Preference Methods for Valuing Environmental Amenities. *Journal of Environmental Economics and Management.* (26):271-292.

Adamowicz, W.L. and J. Swait. 1997. Comments on Scaling and Selecting Compensatory Restoration Projects: An Economic Perspective. Presented at the Conference on Restoration of Lost Human Uses of the Environment. Washington, D.C. May 7.

Adamowicz, W.L., J. Swait, P. Boxall, J. Louviere, M. Williams. 1997a. Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation. Forthcoming, *Journal of Environmental Economics and Management.*

Alberini, A. 1995. Efficiency vs. Bias of Willingness to Pay Estimates: Bivariate and Interval Data Models. *Journal of Environmental Economics and Management.* 29:169-180.

Anderson, N.H. 1962. Application of an Additive Model to Impression Formation. *Science.* 138:817-818.

Anderson, N.H. 1970. Functional Measurement and Psychophysical Judgment. Psychological *Review.* 77:153-170.

Anderson, N.H. 1981. *Foundations of Information Integration Theory*. New York: Academic Press.

Anderson, N.H. 1982. *Methods of Information Integration Theory*. New York: Academic Press.

Anderson, S., A. de Palma and J. Thisse. 1992. *Discrete Choice Theory of Product Differentiation*, Cambridge:MIT Press.

Anderson, N.H. and J. Shanteau. 1977. Weak Inference With Linear Models. *Psychological Bulletin.* 84:1155-1170.

Barron, F.H. 1977. Axiomatic Conjoint Measurement. *Decision Sciences.* 8:548-559.

Batsell, R.R. and J.J. Louviere. 1991. Experimental Analysis of Choice. *Marketing Letters.* 2:99-214.

Ben-Akiva, M. and S.R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application To Travel Demand*. Cambridge, MA: MIT Press.

Ben-Akiva, M. and T. Morikawa. 1991. Estimation of travel demand models from multiple data sources, in M. Koshi (ed.) Transportation and Traffic Theory, Proceedings of the 11th ISTTT, Elsevier, Amsterdam, 461-476.

Ben-Akiva, M., T. Morikawa and F. Shiroishi. 1992. Analysis of the Reliability of Preference Ranking Data. *Journal of Business Research.* 24:149-164.

Blamey, R. 1996. Citizens, Consumers and Contingent Valuation: Clarification and the Expression of Citizen Values and Issue-Opinions. In Adamowicz, W.L., P.C. Boxall, M.K. Luckert, W.E. Phillips and W.A. White. (eds). *Forestry, Economics and the Environment*. CAB International, Oxon, U.K. 103-133.

Boxall, P. W. Adamowicz, M. Williams, J. Swait and J. Louviere. 1996. A comparison of stated preference approaches to the measurement of environmental values. *Ecological Economics.* 18:243-253.

Boyle, K.J., M.P. Welsh and R.C. Bishop. 1993. The Role of Question Order and Respondent Experience in Contingent-Valuation Studies. *Journal of Environmental Economics and Management.* 25:S80-S99.

Brazell, J. and Louviere, J.J. .1997. Length Effects On Response Rates And Reliability. Unpublished Working Paper, Department of Marketing, Faculty of Economics, University of Sydney, January. Previous versions presented at the 1996 Marketing Science Conference, Gainesville, Fl (March 1996).

Brunswick, E. 1952. *The Conceptual Framework of Psychology*. Chicago: University of Chicago Press.

Bunch, D.S. 1991. Estimability in the Multinomial Probit Model. *Transportation.* 25B(1):1-12.

Bunch, D.S. and R.R. Batsell. 1989. How Many Choices Are Enough? The Effect of the Number of Observations on Maximum Likelihood Estimator Performance in the Analysis of Discrete Choice Repeated-Measures Data Sets with the Multinomial Logit Model. Unpublished Working Paper, Graduate School of Management, University of California, Davis.

Bunch, D.S., J. Louviere and D.A. Anderson. 1996. A Comparison of Experimental Design Strategies for Choice-Based Conjoint analysis with Generic-Attribute Multinomial Logit Models. Working Paper, UC Davis Graduate School of Management, May.

Cameron, T. 1992. Combining Contingent Valuation and Travel Cost Data for the Valuation of Non-market Goods. *Land Economics.* 68(3):302-317.

Cameron, T. 1997. Credibility of Information Sources and the Formation of Individuals' Option Prices for Climate Change Mitigation. Paper presented at the *EML/NSF Symposium on Preference Elicitation*, UC Berkeley (July 29-August 1, 1997).

Cameron, T. and J. Englin. 1997. Respondent Experience and Contingent Valuation of Environmental Goods. *Journal of Environmental Economics and Management.* 33:296-313.

Cameron, T. and J. Quiggin. 1994. Estimation Using Contingent Valuation Data from a 'Dichotomous Choice with Follow-up' Questionnaire. *Journal of Environmental Economics and Management.* 27(3):218-234.

Carson, R.T. N. Flores, K. Martin, and J. Wright. 1996. Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods. *Land Economics.* 72(1):80-99.

Carson, R.T. , T. Groves and M.J. Machina. 1997. Stated Preference Questions: Context and Optimal Response. Paper presented at the EML/NSF Symposium on Preference Elicitation, July 29 – August 31, 1997. Berkeley, CA.

Carson, R., J. Louviere, D. Anderson, P. Arabie, D. Bunch, D. Hensher, R. Johnson, W. Kuhfeld, D. Steinberg, J. Swait, H. Timmermans, B. Wiley. 1994a. Experimental Analysis of Choice. Special Issue on the 1993 Duke Choice Symposium, *Marketing Letters.* 5(4):351-368.

Carson, R.T. and R.C. Mitchell. 1995. Sequencing and nesting in contingent valuation surveys. *Journal of Environmental Economics and Management.* 28:155-174.

Carson, R.T., J. Wright, A. Alberini, N. Carson and N. Flores. 1994b. A Bibliography of Contingent Valuation Studies and Papers. Natural Resource Damage Assessment. La Jolla, CA.

Cattin, P. and D.R. Wittink. 1982. Commercial Use of Conjoint Analysis: A Survey. *Journal of Marketing.* 46:44-53.

Chapman, R.G. 1984. An Approach to Estimating Logit Models of a Single Decision Maker's Choice Behavior. *Advances in Consumer Research.* 11:656-661.

Chapman, R.G. and R. Staelin. 1982. Exploiting Rank Ordered Choice Set Data Within The Stochastic Utility Model. *Journal of Marketing Research.* 19:288-301.

Cosslett, S. 1981. Efficient Estimation of Discrete Choice Models. In *Structural Analysis of Discrete Data with Econometric Applications*, C. Manski and D. McFadden, editors, Cambridge: MIT Press, 51-111.

Cox, D.R. 1970. *Analysis of Binary Data*. London: Methuen.

Cummings, R. G., P. T. Ganderton, and T. McGuckin. 1994. Substitution Effects in CVM Values. *American Journal of Agricultural Economics.* 76:205-214.

Daganzo, C. 1980. *Multinomial Probit*. New York: Academic Press.

Dawes, R.M. and B. Corrigan. 1974. Linear Models in Decision Making. *Psychological Bulletin.* 81:95-106.

Desvousges, W.H., E. E. Fries, S.P. Hudson and R.W. Dunford. 1997. Scaling and Selecting Compensatory Restoration Projects: An Economic Perspective. Triangle Economic Research, Durham, NC.

Desvousges, W.H., K.R. Mathews, R.W. Dunford, R.R. Russell and H.L. James. 1995. Comments on NOAA's Proposed NRDA Rule Under the Oil Pollution Act. Triangle Economic Research, Durham, NC.

Domencich, T.G. and D. McFadden. 1975. *Urban Travel Demand: A Behavioral Analysis*. Amsterdam: North-Holland.

Elrod, T.S. and S.K. Kumar. 1989. Bias in the First Choice Rule for Predicting Share. Proceedings of the Sawtooth Software Conference, Ketchum, ID: Sawtooth Software, 259-271.

Elrod, T.S., J.J. Louviere and K.S. Davey. 1992. An Empirical Comparison of Ratings-Based and Choice-Based Conjoint Models. *Journal of Marketing Research.* 29:368-377.

Erdem, T. 1993. Dynamic Brand Choice and Product Marketing Signaling Under Imperfect Information. Unpublished PhD Thesis, Faculty of Business, University of Alberta, Edmonton, Alberta.

Feather, P. 1994. Sampling and Aggregation Issues in Random Utility Model Estimation. *American Journal of Agricultural Economics*. 76(4): 772-80.

Finn, A. and J.J. Louviere. 1992. Determining The Appropriate Response To Evidence of Public Concern: The Case of Food Safety. *Journal of Public Policy and Marketing*. 11(1):12-25.

Finn, A., J.J. Louviere and H. Timmermans, 1992. Retail Research. In *Handbook of Marketing Research*. M. Houston (Ed.). New York: McGraw-Hill.

Freeman, A.M. 1991. Factorial survey methods and willingness to pay for housing characteristics: comment. *Journal of Environmental Economics and Management.* 20:92-96.

Gaudagni, P.M. and J.D.C. Little. 1983. A Logit Model of Brand Choice Calibrated on Scanner Data. *Marketing Science.* 2:203-238.

Gensch, D. 1985. Empirically Testing a Disaggregate Choice Model for Segments. *Journal of Marketing Research*. 22:462-467.

Goodman, A.C., 1989. Identifying willingness-to-pay for heterogeneous goods with factorial survey methods. *Journal of Environmental Economics and Management.* 16:58-79.

Green, P.E. 1974. On the Design of Choice Experiments Involving Multifactor Alternatives. *Journal of Consumer Research*. 1:61-68.

Green, P.E. 1992. On the Non-Use of Contingent Valuation Methodology in Marketing Research. The Wharton School, University of Pennsylvania.

Green, P.E. 1995. The Role and Limitations of Conjoint Analysis in the Preference Modeling of Constructed Markets. The Wharton School, University of Pennsylvania.

Green, P.E., K. Helsen and B. Shandler. 1988. Conjoint Validity Under Alternative Profile Presentations. *Journal of Consumer Research.* 15:392-397.

Green, P.E. and V.R. Rao. 1971. Conjoint Measurement for Quantifying Judgmental Data. *Journal of Marketing Research.* 8:355-363.

Green, P.E. and V. Srinivasan. 1978. Conjoint Analysis in Consumer Research: Issues and Outlook. *Journal of Consumer Research.* 5:103-123.

Green, P.E. and V. Srinivasan. 1990. Conjoint Analysis in Marketing Research: New Developments and Directions. *Journal of Marketing.* 54(4):3-19.

Green, P.E. and Y. Wind. 1973. *Multiattribute Decisions in Marketing: A Measurement Approach*. Hinsdale, IL: Dryden Press.

Gregory, R., S. Lichtenstein, and P. Slovic. 1993. Valuing Environmental Resources: A Constructive Approach. *Journal of Risk and Uncertainty*. 7(2):177-97.

Haggerty, M.R. 1985. Improving the Predictive Power of Conjoint Analysis: The Use of Factor Analysis and Cluster Analysis. *Journal of Marketing Research.* 22:168-184.

Hailu, A., W. Adamowicz and P. Boxall. 1996. Substitution and Valuation: Application of a Multi-Program Environmental Valuation Method. Unpublished paper, Department of Rural Economy, University of Alberta, Edmonton, Alberta.

Hammond, K.R. 1955. Probabilistic Functioning and the Clinical Method. *Psychological Review.* 62:255-262.

Hammond, K.R., C.J. Hursch and F.J. Todd. 1964. Analyzing the Components of Clinical Inference. *Psychological Review.* 71:438-456.

Hanemann, W.M. 1991. Willingness to Pay and Willingness to Accept: How Much can they Differ. *American Economic Review.* 81:635-647.

Hanemann, W.M., J.B. Loomis and B. Kanninen. 1991. Statistical Efficiency of Double-bounded Dichotomous Choice Contingent Valuation. *American Journal of Agricultural Economics*. 73(4):1255-1263.

Hayes, D.J., J.F. Shogren, S.Y. Shin and J.B. Klebenstein. 1995. Valuing Food Safety in Experimental Auction Markets. *American Journal of Agricultural Economics.* 77(1):40-53.

Heckman, J. 1981. Statistical Models for Discrete Panel Data. In *Structural Analysis of Discrete Data with Econometric Applications*, C. Manski and D. McFadden, editors, Cambridge: MIT Press, 114-178.

Heckman, J. 1997. Review of Adamowicz, Louviere and Swait. University of Chicago, October, 1997. 4pp.

Hensher, D.A. 1994. Stated preference analysis of travel choices: the state of practice. *Transportation.* 21:107-133.

Hensher, D.A. and J.J. Louviere. 1983. Identifying Individual Preferences for International Air Travel: An Application of Functional Measurement Theory. *Journal of Transport Economics and Policy.* 17:225-245.

Hensher, D.A., J. Louviere and J. Swait. 1997. Combining Sources of Preference Data. Unpublished Working Paper, Graduate School of Management, University of Sydney, Sydney, Australia, December, 35pp.

Hoehn, J. P. and J. Loomis. 1993. Substitution Effects in the Valuation of Multiple Environmental Programs. *Journal of Environmental Economics and Management.* 25:56-75.

Hoffman, P.J. 1960. The Paramorphic Representation of Clinical Judgment. *Psychological Bulletin.* 57:116-131.

Hoffman, P.J., P. Slovic and L.G. Rorer. 1968. An Analysis-of-Variance Model for the Assessment of Configural Cue Utilization in Clinical Judgment. *Psychological Bulletin.* 69:338-349.

Horowitz, J.L. and J.J. Louviere. 1993. Testing Predicted Probabilities Against Observed Discrete Choices in Probabilistic Discrete Choice Models. *Marketing Science.* 12(3):270-279.

Huber, J. and K. Zwerina. 1996. The Importance of Utility Balance in Efficient Choice Designs. *Journal of Marketing Research.* 33:307-317.

Jennings, S.M. 1992. Wood Quality Management in British Columbia. Unpublished Doctoral Dissertation, Department of Rural Economy, University of Alberta, Edmonton.

Johnson, R.D. 1987. Making judgments when information is missing: Inferences, biases, and framing effects. *Acta Psychologica.* 66:69-82.

Johnson, R.D. 1989. Making decisions with incomplete information: The first complete test of the inference model. In T. Srull (Ed.) *Advances in Consumer Research.* 16:522-528.

Johnson. R.D. and I.P. Levin. 1985. More than meets the eye: The effect of missing information on purchase evaluations. *Journal of Consumer Research*. 12:160-177.

Johnson, E., R. Meyer and S. Ghose. 1989. When Choice Models Fail: Compensatory Models in Negatively Correlated Environments. *Journal of Marketing Research.* 26:255-270.

Jones, C. A. and K. A. Pease. 1997. Restoration-Based Measures of Compensation in Natural Resource Liability Statutes. *Contemporary Economic Policy*. XV:111-122.

Kaciak, E. and J.J. Louviere. 1990. Multiple Correspondence Analysis of Multiple Choice Data. *Journal of Marketing Research*. 27:455-465.

Kahneman, D. and J.L. Knetsch. 1992. Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management*. 22:57-70.

Kahneman, D., J.L. Knetsch and R.H. Thaler. 1990. Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy.* 98:1325-1348.

Keeney, R.L. and H. Raiffa. 1976. *Decisions With Multiple Objectives: Preferences and Value Tradeoffs*. New York: John Wiley.

Knetsch, J.L. 1991. The Endowment Effect and Evidence of Irreversible Indifference Curves. *American Economic Review.* 1277-1284.

Krantz, D.H., R.D. Luce, P. Suppes and A. Tversky. 1969. *Foundations of Measurement*. New York: Academic Press.

Krantz, D.H. and A. Tversky. 1971. Conjoint Measurement Analysis of Composition Rules in Psychology. *Psychological Review.* 78:151-169.

Kuhfeld, W., R. Tobias and M. Garratt. 1994. Efficient Experimental Design with Marketing Research Applications. *Journal of Marketing Researc*h. 31:545-557.

Kunreuther, H. and D. Easterling. 1996. The Role of Compensation in Siting Hazardous Waste Facilities. *Journal of Policy Analysis and Management.* 15(4):601-622.

Lancaster, K. 1966. A New Approach to Consumer Theory. *Journal of Political Economy*. 74:132-157.

Lareau, T.J. and D.A. Rae 1988. Valuing WTP for diesel odor reductions: an application of contingent ranking technique. *Southern Economic Journal*. 728-742.

Layton, D.F. 1996. Rank Ordered, Random Coefficients Multinomial Probit Models for Stated Preference Surveys: An Investigation of Preference Heterogeneity, IIA, and Contingent Ranking Data. Paper presented at the 1996 AERE Workshop, June 2-4, Tahoe City, CA.

Lazari, A. and D.A. Anderson. 1994. Designs of Discrete Choice Set Experiments for Estimating Both Attribute and Availability Cross Effects. *Journal of Marketing Research.* 31:375-383.

Levin, I.P., R. Meyer and J.J. Louviere. 1978. Functional analysis of Mode Choice. *Transportation Research Record.* 673:1-7.

Liebetrau, A.M. 1983. *Measures of Association*. Sage University Papers Series on Quantitative Applications in the Social Sciences. Newbury Park, Ca: Sage Publications, Inc.

Louviere, J.J. 1984. Hierachical Information Integration: A New Approach for the Design and Analysis of Complex Multiattribute Problems. *Advances in Consumer Research.* 11:148-155.

Louviere, J.J. 1987. An Experimental Design Approach to the Development of Conjoint Based Simulation Systems With An Application To Forecasting Future Retirement Choices. In *Behavioral Modeling in Geography and Planning*, R.J.G. Golledge and H.G. Timmermans (Eds). London: Croom-Helm.

Louviere, J.J. 1988a. *Analyzing Individual Decision Making: Metric Conjoint Analysis.* Sage University Series on Quantitative Applications in the Social Sciences, Series No. 67. Newbury Park, Ca: Sage Publications, Inc.

Louviere, J.J. 1988b. Conjoint Analysis Modeling of Stated Preferences: A Review of Theory, Methods, Recent Developments and External Validity. *Journal of Transport Economics and Polic*y. 10:93-119.

Louviere, J.J. 1994. Conjoint Analysis. In R. Bagozzi (Ed.) *Advances in Marketing Research*. Cambridge, MA:Blackwell Publishers.

Louviere, J.J. 1996. Relating Stated Preference Measures and Models to Choices in Real Markets: Calibration of CV Responses. In D.J. Bjornstad and J.R. Kahn editors, *The Contingent Valuation of Environmental Resources: Methodological Issues and Research Needs*. Brookfield, U.S: Edward Elgar.

Louviere, J., D. Anderson, J. Swait and J. Gray-Lee. 1995. Cross-Task and External Validity Comparisons of SP and RP Choice Models. Presented at the 1995 International Conference on Retailing and Consumer Services, Gold Coast, Australia.

Louviere, J.J., A. Finn and H.J.P. Timmermans. 1994. Retail Research Methods. In Houston, M.J. (Ed.), *Handbook of Marketing Research*, McGraw-Hill.

Louviere, J.J., M. Fox and W. Moore. 1993. Cross-Task Validity Comparisons of Stated Preference Models. *Marketing Letters.* 4:205-213.

Louviere, J. and J. Swait. 1996. Searching For Regularities In Choice Processes, Or The Little Constant That Could. Working Paper, Dept. of Marketing, Faculty of Economics, University of Sydney, October 1996, Sydney, Australia, 50 pp.

Louviere, J.J. and H.J.P. Timmermans. 1990. Stated Preference and Choice Models Applied in Recreation Research: A Review. *Leisure Sciences.* 12:9-32.

Louviere, J.J. and G.G. Woodworth. 1983. Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data, *Journal of Marketing Research.* 20:350-367.

Luce, R.D. 1959 *Individual Choice Behavior: A Theoretical Analysis*. New York: John Wiley & Sons.

Luce, R.D. and Suppes. 1965. Preference, Utility and Subjective Probability. In Luce, R.D., Bush, R.R, and Galanter, E. (Eds.), *Handbook of Mathematical Psychology*. New York: John Wiley & Sons, 249-410.

Lynch, J.G. 1985. Uniqueness Issues in the Decompositional Modeling of Multiattribute Overall Evaluations: An Information Perspective. *Journal of Marketing Research.* 22:1-19.

Mackenzie, J. 1993. A comparison of contingent preference models. *American Journal of Agricultural Economics.* 75:593-603.

Manski, C. 1977. The Structure of Random Utility Models. *Theory and Decision*. 8:229-254.

McDaniels, T. 1996. The Structured Value Referendum: Eliciting Preferences for Environmental Policy Alternatives. *Journal of Policy Analysis and Management.* 15(2):227-251.

McDaniels, T., L.J. Axelrod and P. Slovic. 1995. Characterizing Perceptions of Ecological Risk. *Risk Analysis.* 15(5):575-588.

McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, New York: Academic Press, 105-142.

McFadden, D. 1975. On Independence, Structure and Simultaneity in Transportation Demand Analysis. Working Paper No. 7511, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley.

McFadden, D. 1981. Econometric Models of Probabilistic Choice. In *Structural Analysis of Discrete Data with Econometric Applications*, C. Manski and D. McFadden, editors, Cambridge: MIT Press, 198-272.

McFadden, D. 1987. Regression-Based Specification Tests for the Multinomial Logit Model. *Journal of Econometrics.* 34:63-82.

McFadden, D. 1994. Contingent Valuation and Social Choice. *American Journal of Agricultural Economics.* 76:689-708.

McFadden, D. 1996. Why is Natural Resource Damage Assessment so Hard? Hibbard Lecture, 12 April 1996, presented at the Department of Agricultural and Resource Economics, University of Wisconsin, Madison. Available at: http://elsa.berkeley.edu/wp/mcfadden0496/readme.html.

Mitchell, R.C., and R.T. Carson. 1989. *Using surveys to value public goods: the contingent valuation method.* Baltimore:Johns Hopkins University Press for Resources for the Future.

Moore, W.L. and D.R. Lehmann. 1990. A Paired Comparison Nested Logit Model of Individual Preference Structures. *Journal of Marketing Research*. 26:420-428.

Neill, H. R. 1995. The Context for Substitutes in CVM Studies: Some Empirical Observations. *Journal of Environmental Economics and Management.* 29:393-397.

Oliphant, K., T. Eagle, J. Louviere and D.A. Anderson. 1992. Cross-Task Comparison of Ratings-Based and Choice-Based Conjoint. Paper presented to the American Marketing Association's Advanced Research Techniques Forum, Lake Tahoe, June. Also presented to the Joint Sawtooth Software, SYSTAT Conference, Sun Valley, ID, July.

Olsen. G.D. and J. Swait. 1997. Nothing is Important. Working Paper, Advanis Inc., 12 W. University Ave. #205, Gainesville, FL 32601, 28pp.

Olsen, G.D., J. Swait, R.D. Johnson, and J.J. Louviere. 1995. Response Mode Influences on Attribute Weight and Predictive Ability When Linear Models are Not Certain to be Robust. Working paper, Faculty of Management, University of Calgary, Sept. 1995, 30 pp.

Opaluch, J.J., S. Swallow, T. Weaver, C Wessels, and D. Wichlens. 1993. Evaluating impacts from noxious waste facilities: including public preferences in current siting mechanisms. *Journal of Environmental Economics and Management*. 24:41-59.

Parsons, G. and M.J. Kealy. 1992. Randomly Drawn Opportunity Sets in a Random Utility Model of Lake Recreation. *Land Economics.* 68(1):93-106.

Peters, T., W. Adamowicz and P. Boxall. 1995. The Influence of Choice Set considerations in Modelling the Benefits from Improved Water Quality. *Water Resources Research.* 31(7):1781-1787.

Peterson, G., T. Brown, D. McCollum, P. Bell, A. Berjulin, A. Carke. 1995. Moral responsibility effects in valuation of WTA for public and private goods by the method of paired comparison. In *Forestry, Environment and Economics*, W. Adamowicz, P. Boxall, M. Luckert, W. Phillips and W. White (eds.) Oxon, U.K., CAB International..

Piattelli-Palmarini, M. 1994. *Inevitable Illusions: How Mistakes Rule Our Minds*. New York: John Wiley & Sons, Inc.

Rae, D.A. 1983. The value to visitors of improving visibility at Mesa Verde and Great Smokey National Parks, in Rowe, R.D. and L.G. Chestnut (eds), *Managing Air Quality* *and Scenic Resources at National Parks and Wilderness Areas.* Boulder, CO:Westview Press.

Slovic, P. and S. Lichtenstein. 1971. Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment. *Organizational behavior and Human Performance.* 6:649-744.

Small, K. 1981. A Discrete Choice Model for Ordered Alternatives. *Econometrica.* 55(2):409-424.

Swait, J. 1984. Incorporating Constraints in Discrete Choice Models of Transportation Demand. Unpublished PhD Thesis, Dept. of Civil Engineering, Massachusetts Institute of Technology.

Swait, J. 1994. A Structural Equation Model of Latent Segmentation and Product Choice for Cross-Sectional, Revealed Preference Choice Data. *J. of Retailing and Consumer Services.* 1(2):77-89.

Swait, J. and W. Adamowicz. 1997. The Effect of Choice Environment and Task Demands on Consumer Behavior: Discriminating Between Contribution and Confusion. Working Paper, Dept. of Rural Economy, Fac. Of Agriculture, Forestry and Home Economics, University of Alberta, Edmonton, February, 39 pp.

Swait, J. and M. Ben-Akiva. 1985. An Analysis of the Effects of Captivity on Travel Time and Cost Elasticities. Annals of the 1985 International Conference on Travel Behavior, April 16-19, Noordwijk, Holland, 113-128.

Swait, J. and M. Ben-Akiva. 1986. Constraints on Individual Travel Behavior in a Brazilian City. *Transportation Research Record.* 1085:75-85.

Swait, J. and M. Ben-Akiva. 1987a. Incorporating Random Constraints in Discrete Models of Choice Set Generation. *Transportation Research B.* 21B(2):91-102.

Swait, J. and M. Ben-Akiva. 1987b. Empirical Test of a Constrained Choice Discrete Model: Mode Choice in São Paulo. *Transportation Research B.* 21B(2):103-115.

Swait, J., T. Erdem, J. Louviere and C. Dubelaar. 1993. The Equalization Price: A Measure of Consumer-Perceived Brand Equity. *International Journal of Research in Marketing.* 10:23-45.

Swait, J. and J. Louviere. 1993. The Role of the Scale Parameter in The Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing Research*. 30:305-314.

Swait, J., J. Louviere and M. Williams. 1994. A Sequential Approach to Exploiting the Combined Strengths of SP and RP Data: Application to Freight Shipper Choice. *Transportation.* 21:135-152.

Thurstone, L.L. 1927. A Law of Comparative Judgment. *Psychological Review.* 4:273-286.

Timmermans, H.J.P. and Heijden, R. van der. 1987. Uncovering Spatial Decision Making Processes: A Decision Net Approach Applied To Recreational Choice Behaviour. *Tijdschrift voor Economische en Sociale Geografie.* 78:298-305.

Train, K., D. McFadden and M. Ben-Akiva. 1987. The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Charges. *Rand Journal of Economics.* 18(1):109-123.

Viscusi, W.K., W.A. Magat, and J. Huber. 1991. Pricing environmental health risks: survey assessments of risk-risk and risk-dollar tradeoffs for chronic bronchitis. *Journal of Environmental Economics and Management.* 21:32-51.

Wainer, H. 1976. Estimating Coefficients in Linear Models: It Don't Make No Nevermind. *Psychological Bulletin.* 83:213-217.

White, H. 1982. Maximum Likelihood Estimation of Misspecified Models. *Econometrica.* 50(1):1-25.

Wittink, D.R. and P. Cattin. 1989. Commercial Use Of Conjoint Analysis: An Update. *Journal of Marketing.* 53:91-98.

Yellott, J.I. 1977. The Relationship Between Luce's Choice Axiom, Thurstone's Theory of Comparative Judgment, and the Double Exponential Distribution. *Journal of Mathematical Psychology.* 15:109-144.

Zahoric, A. and V.R. Rao. 1990. An Application of the Beta-Distribution in Aggregate Conjoint Analysis. Paper presented to the TIMS Marketing Science Conference, Durham, NC, March.