MODELLING OF UNINTENDED EFFECTS

WITH PANEL INFORMATION IN

STATED PREFERENCE NON-MARKET VALUATION:

APPROACHES AND APPLICATIONS

by

DMITRIY VOLINSKIY

(Under the direction of John C. Bergstrom)

ABSTRACT

The dissertation is concerned with practical modelling issues in stated preference non-market valuation applications when undesirable (unintended) effects are present. Effects that are common to valuation studies are described and classified. Practical approaches and methods to build empirical models in view of unintended effects are reviewed and discussed. Three case studies follow up to illustrate the range of possibilities in empirical treatment of unintended effects. Case Study I implements a panel mixed logit model to account for respondents' heterogeneity in decision-making rules, a sample selection bias, and "warm glow." Case Study II is dedicated to respondents' rationality in a sequential multiple commodity valuation setting. It develops a stochastic model which is consistent with the axioms of reflexivity, transitivity, and continuity. Case Study III presents a fuzzy logic system as a practical solution to the issue of impossibility of exactly quantifying qualitative categories in commodity description and elicitation format options.

INDEX WORDS: Stated preferences, Contingent valuation, Biases, Mixed logit,
                Sequential choice, Fuzzy expert systems

MODELLING OF UNINTENDED EFFECTS

WITH PANEL INFORMATION IN

STATED PREFERENCE NON-MARKET VALUATION:

APPROACHES AND APPLICATIONS

by

DMITRIY VOLINSKIY

Engineer, Tashkent State Technical University, Uzbekistan, 1995

M.S., Louisiana State University, 2001

A Dissertation Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2005

MODELLING OF UNINTENDED EFFECTS

WITH PANEL INFORMATION IN

STATED PREFERENCE NON-MARKET VALUATION:

APPROACHES AND APPLICATIONS

by

DMITRIY VOLINSKIY

Approved:

Major Professor: John C. Bergstrom

Committee: Christopher M. Cornwell
Dmitry Vedenov

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2005

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND AND OBJECTIVES

Stated-preference valuation methods of non-market commodities such as environmental amenities, programs, or policies have a unique place not only in environmental economics but, perhaps, in economics in general. Stated-preference valuation is truly interdisciplinary. Primarily an economics exercise, it nonetheless embraces decision theory, behavioral psychology, sociology, political science, as well as many other fields and disciplines.

Every field of knowledge involved has concerns of its own. While the economic investigator would aim to measure a welfare change emanating from the provision or ceasing to provide the non-market commodity, other disciplines would be likely to focus on non-economic phenomena that occur in stated-preference valuation experiments. A decision theorist may be interested in a set of rules agents employ in making decisions with respect to the commodity in question. A psychologist's interest may be how commodity information is perceived and what motivates the agent to behave in a particular way. A sociologist's attention might be drawn to the role of demographic factors and so on.

As a result of this large interdisciplinary involvement, numerous phenomena have been described as potential pitfalls in stated-preference valuation experiments. Economics does not provide a treatment for most of them. For instance, those phenomena cannot reasonably enter the traditional cost-benefit analysis that the applied stated-preference study may be part of. In other words, these occurrences are not necessarily nuisances but, at the same time, they are not intended to be part to the economic model. They are *unintended effects* from the viewpoint of applied economic research.

Meanwhile, these unintended effects do take place and need to be dealt with. Demand theory is deductive. A model of economic behavior for the representative agent is an abstraction to explain observed trends in the behavior of large masses of people operating out in the market. If the law of large numbers can apply socially, as some argue, then averaging over legions of agents should diminish the prominence of non-economic effects, leaving the pure self-interest as the driving force behind economic choices. Stated-preference valuation, on the contrary, is inductive. There is no market yet, hence no representative agent. Extrapolation on the basis of survey information is required. Survey samples in valuation experiments are small and respondents reaction is generally contingent upon the survey setting. Regarding the survey sample as highly representative and informative for extrapolation purposes may lead to the gambler's fallacy (Tversky and Kahneman 1971). Chances are that, instead of a deviation in one direction being cancelled by a deviation in the opposite direction, effects may be co-directional. Matching magnitudes for the cancelling may also be problematic with small samples.

As a result, obtained welfare change measure may be biased, unreliable, and/or impure. Biasedness here refers to under- or overestimating the welfare change. Unreliability means a large spread of possible values of the measure, rendering it unusable for policy-making purposes. Impurity refers to the composite nature of measure estimates, which does not allow to separate the desired economic effect from unintended ones. Either of these qualities threatens to invalidate the results of the stated-preference study.

Modelling of unintended effects is the use of mathematical representation and methods in order to protect and enhance the quality of information to be delivered by the stated-preference valuation study. Three general directions in the modelling can be identified, ranked by their conservativeness: (a) detecting and removing observational units that feature a particular effect, (b) building a robust model, and (c) making observational units that feature the effect an informative part of the model.

Modelling of unintended effects in stated-preference valuation has received considerable attention in scholarly literature. Most studies, however, target a particular effect or class of effects. Many effects have been described, sometimes under different names, which creates an impression that a valuation experiment is a decision-making paradox. The first major objective of this dissertation is to provide a review and classification of unintended effects alongside approaches and methods to model them. Two seemingly different effects can be represented in a similar way, provided that they share a common genesis. This would help an applied investigator narrow the circle of mathematical methods and techniques to deal with a suspected effect. The classification of effects *per se* is deemed to be of importance, since it is instrumental in revealing linkages between various effects, thus tracing their precursors and consequences. This constitutes one part of the dissertation: the approaches.

Modern technology such as computer-aided surveys and the Internet provides the facilities to conduct sophisticated valuation experiments and obtain a large array of information from a single respondent. Different commodities, different levels of the same commodity, or different bid amounts can be sequentially offered to the respondent. Panel data structures are thus generated. These provide the investigator with richer information; these also bring about new effects and new concerns.

With all these advances, the modelling toolkit of the valuation practitioner remains quite limited. There is an evident gap between theoretical developments on one hand, and practice, on the other. This leads to the second major objective of the study. It is to adapt or extent empirical models from economics, as well as other disciplines, to the benefit of practical modelling of unintended effects in environmental valuation. Three empirical Case Studies included demonstrate a broad range of possibilities for an applied environmental economist in the technology of model building and estimation in the presence of unintended effects. Meeting this objective is the other part of the study: the applications.

The present dissertation is not meant to be a comprehensive review or critique of stated-preference valuation methods in general. Their design and validity have been and will remain

the subject of a never-ending debate. Nor does it aim to investigate diverse phenomena surrounding stated preferences. The dissertation has a clear applied orientation. In what follows, Chapter 2 contains a review and classification of unintended effects as well as approaches to representing them in a mathematical model. Chapters 3–5 are empirical Case Studies. Chapter 3 presents a static random coefficient panel model to estimate willingness-to-pay for farmland conservation easement programs in Georgia. The model accounts for heterogeneity of respondents, selection issues, and the "warm glow" effect. Chapter 4 considers a seemingly dynamic phenomenon. The study in Chapter 4 proposes a discrete-choice model for environmental policy or program valuation, to be used in cases when several alternatives are valued sequentially. Chapter 5 is dedicated to the heterogeneity in perception. It looks at a contingent valuation experiment as a problem of mapping entities that are not exactly quantifiable. The use of a fuzzy logic advisor for people's judgments is considered and illustrated for a case of Southern Pine Beetle damage restoration. Finally, Chapter 6 concludes with recommended directions for future research.

1.2   Notation and Conventions

The author adheres to standard notation of economic and statistical literature throughout the study. Special notation and conventions are explained below.

Symbol	Denotation
$\mathbf{a}$	vector (boldface, lowercase Latin)
$\mathbf{A}$	matrix (boldface, uppercase Latin)
FACTOR	factor variable (small capitals)
$\{\ldots\}$	collection of sets or sequence
$\langle \mathbf{a}_1 \vert \mathbf{a}_2 \rangle$	inner product of $\mathbf{a}_1$ and $\mathbf{a}_2$
$\Vert \mathbf{a} \Vert$	supremum norm of $\mathbf{a}$
$\mathbf{A}_1 * \mathbf{A}_2$	element-by-element (Hadamard) product of matrices $\mathbf{A}_1$ and $\mathbf{A}_2$

2.1 Decision-Making with Stated Preferences

"Ask a question and you will get an answer," coined Arrow (Arrow and Fisher 1974) summarizing the stated-preference valuation experiment. A valuation practitioner often sees the experiment this way. She would focus on its very beginning, choosing what questions to ask and how to ask them, and the very end, deciding how to use the answers to infer on a selected welfare change measure. Meanwhile, unintended effects occur throughout the decision-making system of the respondent. It is important therefore first to lay out a structure of this system, in order to pinpoint locations of effects and understand their similarities and differences.

Figure 2.1 presents a very general, admittedly technocratic scheme of the respondent's decision-making process, derived from consumer choice models in consumer behavior literature (Blackwell, Miniard, and Engel 2001) and the idea of a "perceptual computer" (Mendel 2001; Zadeh 1999). The scheme has four major components through which information flows and by which it is changed and re-combined.

The first, input unit called the **Cognition Unit** is responsible for receiving commodity and other relevant information from the survey instrument and representing it in a form understandable and interpretable to the respondent. Two streams of information are received and processed by the Cognition Unit. The first is Commodity Information as presented in the survey instrument. If the commodity is indivisible and gestalt, only a qualitative description is available, i.e. no attributes are singled out, then the Cognition Unit processes the single stream. If the commodity is divisible and therefore its quantity is available, the

Figure 2.1: Decision-Making Structure of Stated-Preference Experiments
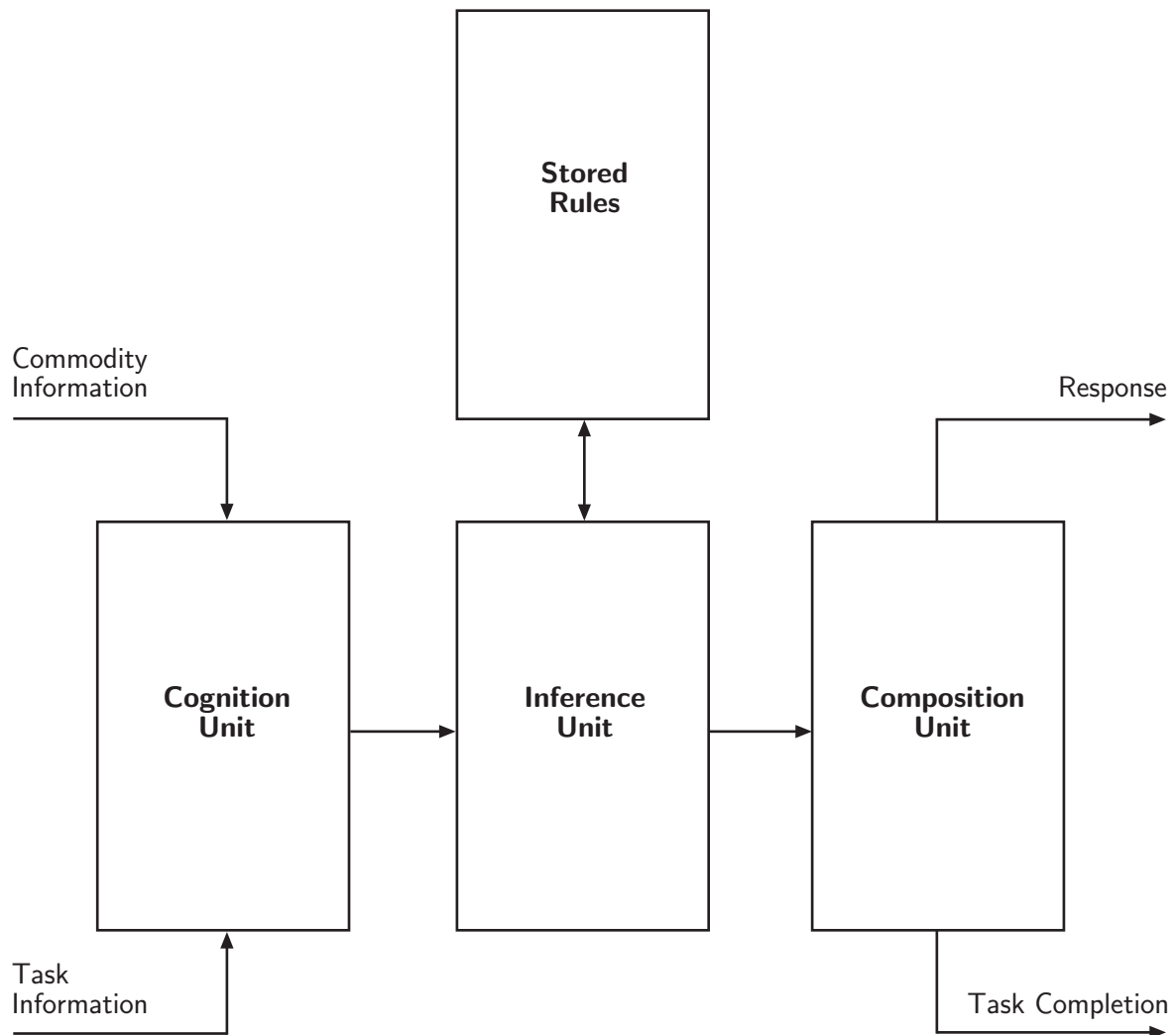
stream comprises the description and the numerical or categorical quantity. Lancaster (1966) suggested a theory whereby these are attributes of the commodity that determine the utility it provides, and therefore utility derived from any particular commodity can be expressed as a function of the generic attributes. If commodity information is supplied broken down by attributes, the respective attribute information streams are processed. Finally, if several commodities are presented at once, the stream conveys the whole choice set information. The second stream — Task Information — carries information about the valuation scenario, which includes a description of the payment vehicle. This stream informs the respondent of what she is expected to do about the commodity in question. The Task Information stream also includes the monetary bid or the "stated cost" of supplying the commodity, if the former is present in the survey instrument. As the author believes, the monetary bid should not be regarded as an attribute of the commodity's. No matter what cost might be stated, it does not alter the commodity itself. It changes the provision scenario, therefore it should be part thereof.

The role of the **Inference Unit** is to handle the way in which **Stored Rules** are activated and processed. The Inference Unit performs three functions. First, depending on the input from the Cognition Unit, it searches for a rule or rules to be applied to commodity information in light of the supplied task. The rule or rules should tell the respondent about how to transform (compose) the received information into a reaction. Once the rule or rules are found, the Inference Unit uses them to carry out the composition. After the composition is complete, the raw output signal is passed on to the output of the system. Finally, feedback from the composition can be used to construct a new rule or rules. If such are generated, then the third function of the Inference Engine is to send the new data back to the Stored Rules unit.

What are the rules the decision-maker uses? These are a collection of IF-THEN statements. These can be represented as functions, classes of functions, or literally, as discrete

statements. No matter what way there are represented in a mathematical model, rules are the "cause-consequence" guides in one's decision-making.

The last unit, the **Composition Unit** performs the reverse of the Cognition Unit's function. Based on information in the task signal, it converts the raw reaction from the Inference Unit into a representation requested by the elicitation format. The processed output can be a number or the "yes/no" answer, or a pointer to a particular choice or choices in the choice set. Put simply, the Composition Unit produces the Response signal. Apart from the Response signal, however, another signal is generated but not necessarily read from this unit. This is a Task Completion signal. It characterizes the level of comfort or satisfaction with the response produced, and it is an important part of the total output. If this signal is not read, i.e. if there is no channel provided for it in the survey instrument, the Task Completion signal is mixed with the Response signal and the mix is supplied to the investigator.

Let me briefly illustrate the working of the decision-making system in Figure 2.1 on an example of the textbook representative consumer. Economic consumer theory states that consumers are rational decision-makers. That is, when faced with a set of possible consumption commodities/bundles, consumers determine their preferences of each of the various bundles and then choose the most preferred commodity/bundle from the set of affordable alternatives. Utility theory axioms therefore shape the design of the Inference Unit and Stored Rules units. The latter keeps a single rule, which is the consumer's indirect utility function. The Cognition Unit simply relays the commodity information and stated cost to the Inference Unit. The Inference engine produces the preference response by applying the indirect utility function rule. The Composition Unit formats the utility levels according to the suggested elicitation protocol and provides the output response.

With the above decision-making structure available, the next step is to review unintended effects and place them in the structure. This is done in Section 2.2.

2.2 Review and Classification of Unintended Effects

An expert panel, headed by Nobel laureates Kenneth Arrow and Robert Solow, summarized the state of the art of stated-preference valuation research in their 1993 "Report of the National Oceanic and Atmospheric Administration Panel on Contingent Valuation" (Arrow et al. 1993), further referred to as the NOAA Report. The NOAA Report points out the risks in valuation surveys, identifies important techniques for avoiding those pitfalls, and ultimately concludes that proper surveys can generate important information about valuation. The report is dedicated specifically to the contingent valuation method. However, contingent valuation epitomizes stated-preference valuation in general. Other methods such as conjoint analysis or iterative bidding games share many characteristics and concerns with contingent valuation, therefore the report can be viewed as one addressing stated-preference valuation methods in general.

Section II of the NOAA Report describes a number of drawbacks to the contingent valuation technique, as deemed important by the panel. Section IV presents guidelines for a valuation study to adhere, should the study produce reliable and thus useful information. These two sections, let alone the rest of the document, are the prime source of information on unintended effects that are of both theorist's and practitioner's concern.

The NOAA Report raises the issue of rational choice in the respondent's behavior. It notes:

> Some of the empirical results produced by CV studies have been alleged to be inconsistent with the assumptions of rational choice... Rationality in its weakest form requires certain kinds of consistency among choices made by individuals... It is reasonable to suppose that more of something regarded as good is better so long as an individual is not satiated. This is in general translated into a willingness to pay [WTP] somewhat more for more of a good, as judged by the individual. Also, if marginal or incremental willingness to pay for additional amounts does decline with the amount already available, it is usually not reasonable to assume that it declines very abruptly.

These passages are related to a phenomenon dubbed the "embedding effect" or "insensitivity to scope."[1] The effect consists in seeming satiation of the consumer with the smallest quantity of the commodity offered in the valuation survey. Several studies demonstrate how WTP may not increase with an increased quantity of the good. (Kahneman 1986) found that WTP for cleaning all lakes in Ontario was insignificantly higher than WTP for cleaning up lakes in just one region. Desvouges et al. (1993) and Schkade and Payne (1994) each found that respondents' WTP to protect migratory birds from dying because of oil spills did not change much whereas the intervention described went up from saving 2,000 birds to 200,000 birds. Commenting on an earlier valuation study, Diamond and Hausman (1993) noted that people's WTP for better visibility in the entire Northwest was hardly any higher than WTP for better visibility in the Grand Canyon, alone. An extensive literature review on suspected cases of the embedding effect can be found in Frederick and Fischhoff (1997).

Two reasons of why embedding takes place are prevalent in economic literature (Frederick and Fischhoff 1997; Barron and Greene 1996). These are the respondent's lack of familiarity and/or experience with the particular non-market commodity and the "warm glow" or purchasing of the do-gooder's feeling, which will be revisited shortly in this section[2]. In terms of placing the embedding effect in the decision-making structure from the previous section, both origins point at the Cognition Unit. The lack of familiarity points at a possibility of the Commodity Information signal being transformed in such a way that its internal representation only allows invoking a very general, top-hierarchy (Green and Tunstall 1999) rule

---

[1]Some studies, e.g. Barron and Greene (1996) and Hanemann (1994), refer to another two phenomena as the embedding effect. These are sub-additivity — the sum of WTP for two different commodities often exceeds the WTP for their combination — and the ordering effect when the commodity that comes first in the survey receives a greater WTP. The sub-additivity is not addressed in this study because, as the author believes, the substitution effect makes it quite natural for related goods and thus there is no unintended effect involved. The ordering effect is treated later in the study as a form of framing.

[2]Barron and Greene (1996) provide evidence in support of another potential cause that they term the "prominence." Prominence consists in people's over-attending the type of the commodity and under-attending its quantity. For the purpose of this study, however, prominence well classifies as a derivative of the lack of experience with the commodity. Respondents cannot well identify the particular commodity and thus fall back to assessing generic types.

such as "if I am asked to vote for or against environmental protection, then …." The possibility of warm glow points on poor processing of the Task Information signal, which totally removes the requested selfishness and egoism in assessing private benefits from the commodity. Accordingly, the embedding effect/scope insensitivity will be classified as belonging to either stream in the Cognition Unit.

The NOAA Report also reflects on a general concern that respondents in stated-preference valuation experiments tend to overstate their WTP ("hypothetical bias"). Again, some respondents inflate WTP to experience the so-called warm glow; that is, satisfaction from being "good citizens." In particular:

> … individuals' responses [may] serve the same function as charitable contributions—not only to support the organization in question, but also to feel the warm glow that attends donating to worthy causes. If this is so, responses should not be taken as reliable estimates of true willingness to pay, but rather as indicative of approval for the environmental program in question … The survey should be designed to deflect the general warm-glow of giving or the dislike of "big business" away from the specific environmental program that is being evaluated.

Warm glow can take form of a contribution; that is, respondents may think of the valuation task as a contribution. It may also emanate from the very fact of participating in a socially desirable undertaking. Regardless of its manifestation, warm glow distorts the respondent's understanding of what her task is and thus firmly classifies as an effect in the Cognition Unit, corrupting the Task Information signal.

Consider another passage from the NOAA Report:

> … a respondent actually willing to pay the stated amount might answer in the negative for several reasons: (i) belief that the proposed scenarios distributed the burden unfairly; (ii) doubt of either the feasibility of the proposed action, so that any contribution would be wasted, or the ability of the relevant agency to carry out the action

efficiently; or (iii) refusal to accept the hypothetical choice problem, because of either a generalized aversion to taxes or a view that someone else ... should pay ...

This effect is known as getting "protest votes." One can notice that protest voting and the warm glow have a great deal in common in spite of their polar consequences. A warm glow effect leads essentially to reporting implausibly high WTP ("yea-saying"), while protest voting reports zero WTP ("nay-saying"). But in both cases the respondent actually pursues her moral, ethical, or political agenda through the valuation instrument. This provides the grounds to conjecture that the warm glow effect is directional, and that its action in the opposite direction gets revealed via protest voting.

The necessity to obtain a large, heterogenous sample of respondents is reiterated throughout the NOAA Report:

Probability sampling is essential for a survey ... Reduction of the final sample by elimination of "protest zeros," "unrealistic high values," and other problematic responses may lead to effective final total response rates so low as to imply that the survey population consists of interested and specially instructed quasi-experts.

Representativeness of a sample requires without any doubt the latter being diverse and heterogenous. Heterogeneity, however, has two sides. Results of a valuation study are likely to be extrapolated onto a larger market. Extrapolation requires some kind of a representative agent. As was mentioned earlier in Chapter 1, a model of economic behavior for the representative agent is an abstraction to explain observed trends in the behavior of large masses of people operating in an actual market. A sample of respondents is no market. It has no cohesive force of a market; correspondingly, the more heterogenous the sample is, the less commonality is left to describe the representative agent. Heterogeneity means that all components of the decision-making system of any given respondent will be different to some degree from those of others. Words mean different things to different people; consequently, one can expect differences in cognition[3]. Differences in the ability to perceive and

---

[3]The NOAA Report provides an example when "forbidding" some action generates less agreement than "not allowing" it, despite the two wordings being logically equivalent.

process commodity information (Bergstrom, Stoll, and Randall 1990) and the information overload (Bergstrom and Stoll 1990), i.e. inability of some respondents to process the scope of commodity information can be examples of heterogeneity in perception. Rules will be different at least in some of their particularities. Socio-economic and demographic characteristics of respondents can affect parameter values in their decision rules. Finally, responses to the elicitation format will be different. As a result, the unintended effects of heterogeneity blanket the entire decision-making system and heterogeneity cannot be restricted to a single component.

The above excerpt also addresses the problem of survey-related non-response. The latter may have two constituents. Sample non-response takes place when people decline the offer of participating in the survey or exit the interview prematurely. Item non-response occurs when respondents do not provide answers to a valuation question. Survey non-response is a large separate problem common to many disciplines and it is beyond the scope of this study. Meanwhile, item non-response has the same distributed nature as heterogeneity. The respondent may fail to produce an answer if she cannot process commodity information or, less likely, understand her task. A rule to apply may be missing. Finally, the respondent may fail to transform her reaction to comply with the elicitation format. Unless there is an indication of a decision-making stage at which a respondent is likely to abort the judgmental process, the safest way would be placing it over the entire system.

As any other social survey, a stated-preference survey is likely to experience some effect of framing. The notion of framing is broad. In general, it leads to a strategic behavior[4] of the respondent induced by the way questions are worded, the survey instrument is set up, or the bare fact of participating in the survey. The NOAA Report notes:

> [speaking of interviewer effects] ... It is possible that interviewers contribute to "social desirability" bias, since preserving the environment is widely viewed as something pos-

---

[4]Hence the term sometimes used for this bias, the "strategic bias."

itive ... the respondent [may] believe that the interviewer would herself favor a yes answer.

The passage above touches upon framing effects of wordings of questions and also what is known in psychology as the "compliance bias." A compliance bias occurs when the respondent attempts to guess the intention of the investigator or interviewer and accommodate it by behaving accordingly (see Keith and Fawson (1996) for empirical evidence of this bias from a contingent valuation study). DeShazo (2002) provides an example of a framing effect manifested as an ordering effect, that is, an apparent change in reported WTP caused by a change in order in which several commodities or monetary bids are presented to the respondent. In his example, two bids, $10 and $5 are offered in the ascending sequence for the same commodity in a bidding game experiment. The higher bid of $10 gets negatively framed by the respondent if she accepts $5. The reason of the negative reaction towards the higher bid is her feeling of loss. Expecting bargaining, the respondent feels she could have rejected the bid of $5 and then have agreed on a smaller amount to follow, buying the commodity cheap. This example also illustrates anchoring, which happens when the respondent uses the first encountered commodity or bid as the point of reference for subsequent choices.

Common to the framing effects, there is a process of adding an imaginary component to commodity and/or task information. Put otherwise, the respondent's inferential process generates more information than was provided. The variety of framing effects therefore belongs in the Inference Unit. Curiously, one of the framing effect types, the compliance bias, appears to have its obverse, similar to the warm glow and protest voting. It is mentioned in the NOAA Report as a "heuristic" thinking:

Respondents [may] rely on a set of heuristics ("these environmental accidents [which are priced through the valuation experiment] are seldom as bad as we're led to believe," or "authorities almost always put too good a face on these things"), in effect they will be answering a different question from that being asked.

With the compliance bias, the respondent becomes overtly cooperative with the investigator, trying to foresee her intentions and oblige with "yea-saying." The heuristically generated disbelief signals an unreasonable wariness in an apprehension about being manipulated or deceived. Unlike the warm glow and protest voting though, the compliance bias and this sort of heuristics do not have polar net results.

The NOAA Report considers a splitting of the binary "yes/no" response into several categories. In particular, the "don't know" option is mentioned:

> ... in national surveys close to a quarter of the population will choose the "don't know" response to most attitude questions if it is explicitly offered; yet these same people will select a substantive alternative if "don't know" is not specifically provided ...

Effects of the number, meanings, and sometimes an order of response options on the response message and accordingly value estimates are known as elicitation format effects (Cummings, Brookshire, and Schulze 1986). Elicitation format effects are similar to cognition effects, since both types involve interpretation of survey conditions such as task and output protocol requirements. However, possibilities of internal representation of an incoming signal are richer by far than possibilities for converting the output decision signal to the rigid elicitation format. Hence the problem is more of incompatibility rather than pure interpretation. For example, if the question is posed as "Would you prefer A to B [yes/no]" and the decision rule activated yields "I like A a bit more than B," then the "yes" response will not be fully compatible with the elicitation format, if the respondent understands "yes" as total, unequivocal agreement.

The suggested decision-making structure allows for representation of elicitation format effects in the Composition Unit through the Task Completion signal. In the previous example, the respondent would signal her low satisfaction with the provided response. If no channel is provided for the signal, that is, if the response format is binary, then the signal will be mixed in the response — the investigator will obtain an unreliable "yes" vote.

Figure 2.2: Decision-Making Structure with Effects Placed
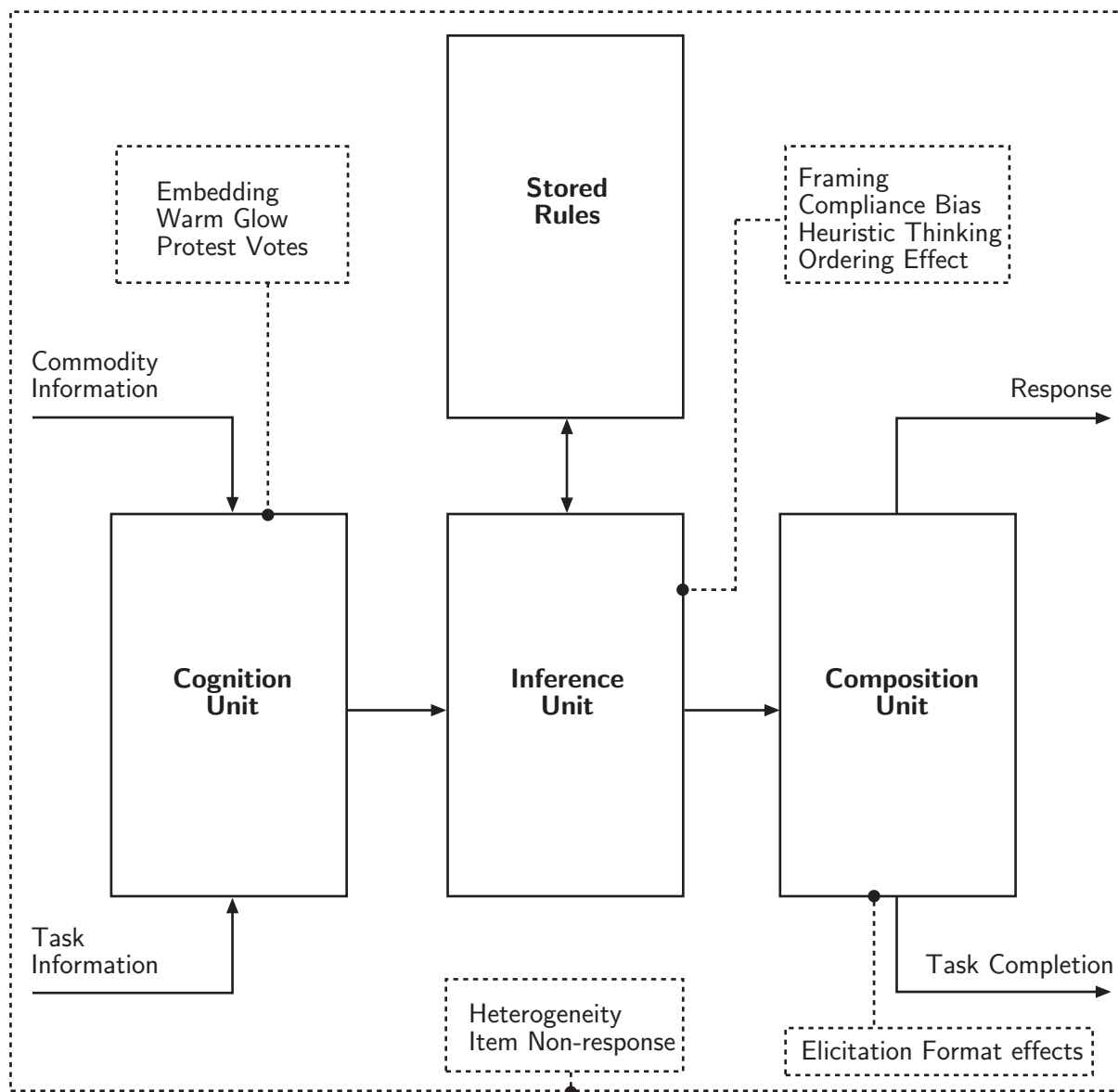
Concluding the description and classification of unintended effects comes Figure 2.2. This is the same decision-making scheme as Figure 2.1, but it now displays the placement of all the effects described in this section. The next section provides a review of approaches to the representation and modelling of these effects.

2.3 Modelling of Unintended Effects

As already mentioned, modelling of unintended effects encompasses three general approaches: (a) detecting and removing observational units that feature a particular effect, (b) building a model robust to the effect, and (c) making observational units that feature the effect an informative part of the model. Methods under (a) are the simplest to implement but lead to a reduction in the sample size. Methods under (b) do not decrease the sample size but may affect the precision of value estimates. Finally, methods under (c) provide the minimum possible loss of estimate quality but may be hard to find.

Insensitivity of WTP estimates to commodity quantity is generally easy to detect, observing statistically insignificant estimates of those parameters in the valuation model that are related to the commodity size or quantity. The use of Lancaster's (1966) theory of commodities as bundles of attributes can help when dealing with an embedding effect/scope insensitivity issue[5]. According to his theory, a commodity itself does not give utility to its consumer. These are characteristics of the commodity's, its attributes, that give rise to utility. This theory is in agreement with the modern perspective on the environmental good valuation problem. An environmental good provides services to the consumer; a change in the services leads to a change in economic values (Bergstrom and Loomis 1999). Such services are mainly qualitative; that is, a service is either present or absent in the bundle of the commodity. Alternatively, several levels or categories of a service can be identified.

---

[5]The author adheres to the equivalence of the notions of embedding and scope insensitivity. An alternative rendition of embedding (Kahneman and Knetsch 1992) as WTP dependence on whether the good is assessed on its own or embedded as part of a more inclusive package has another meaning and therefore can be modelled as an ordinary economic phenomenon (Randall and Hoehn 1996).

Recalling the example from the previous section of scope insensitivity with the protection of migratory waterfowl, protecting 2,000 birds or protecting 200,000 may just mean the overall presence of protection to the consumer. Wordings in that study (Desvouges et al. 1993) even suggested putting both quantities into the same qualitative category, since the quantities were described as a small percentage of the total bird species population.

Hanemann and Kanninen (1999) discuss a multiple commodity valuation model which entirely consists of attribute indicators. With slight changes, it can be written as

$$v = \alpha \; income + \beta \; bid + \sum_i \gamma_i x_i + \sum_{i,j} \gamma_{ij} x_i x_j \qquad (2.1)$$

where $x_i = 1[\text{if } attribute_i \text{ is present}]$; that is, x_i indicates the presence of $attribute_i$. A linear form of this model (all γ_{ij} are set to zero) is used in Case Study II (Chapter 4) of this dissertation for a case of valuing river ecosystem restoration. Management programs in this study include restoring 2, 4, or 6 mile stretches of a river. Correspondingly, there are three indicator variables in the model, one for each quantity. Estimation results indicate the presence of local scope insensitivity: WTP does not increase until the size reaches 6 miles.

The use of qualitatively defined categories on the quantitative commodity size and/or attributes helps remove an embedding. For instance, if, of two integers q_1 and q_2, q_1 is much greater than q_2, then the quantity q_1 of the commodity naturally includes its arbitrary q_2 units. However, a "large quantity" and "small quantity" are different, mutually exclusive categories. This is further illustrated with Venn diagrams in Figure 2.3.

Representing commodity information with qualitative categories can be helpful in modelling the transformation of the Commodity Information signal by the Cognition Unit. Channelling the Response signal in the Composition Unit through a set of qualitative response options can be helpful when dealing with elicitation format effects. With such effects present, the respondent mixes her uncertainty about, or dissatisfaction with the binary "yes/no" response in the response itself, which makes the latter inexact.

Most studies that addressed the inexactness problem in the contingent valuation method were concerned with the respondent's uncertainty about her preferences and, as a result,

Figure 2.3: Qualitative and Quantitative Categories

Numerically

Qualitatively
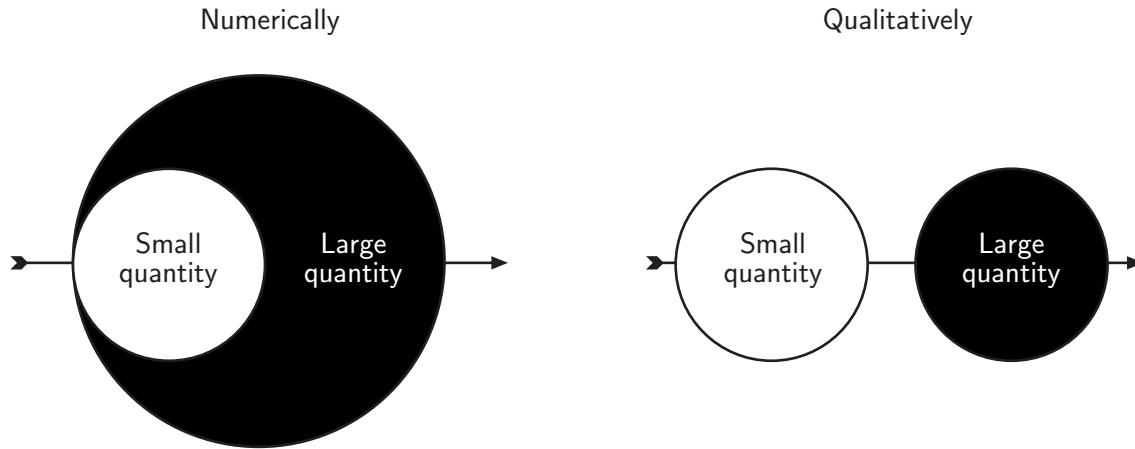
her WTP. In their respective papers, Ready, Navrud, and Dubourg (2001), Wang (1997), Li and Mattsson (1995), and Ready, Whitehead, and Blomquist (1995), suggested ways of partitioning the conventional dichotomous choice of the closed-ended valuation format to make it more continuous.

Li and Mattsson used a follow-up to the dichotomous choice valuation question. Their follow-up question was posed as: "How certain were you of your answer to the previous question?"; the follow-up response was elicited on a scale from 0% (total uncertainty) to 100% (complete certainty). Wang suggested a treatment of the "don't know" option as a latent "yes" or "no." Ready, Whitehead, and Blomquist used a polychotomous choice format to elicit the probability of the "yes" to the valuation question and grades of preference intensity. The former partitioning took form of "definitely yes," "maybe yes," etc. statements, while the latter suggested respondents to use wordings like "strongly prefer," "slightly prefer," and so on. A similar approach was taken in the more recent study by Ready, Navrud, and Dubourg, which supplemented the dichotomous choice format with a payment card experiment.

While multiple categories provide a more flexible structure, any discrete-choice model is still based on the assumption that the unobservable quantity of interest (such as WTP) falls

Figure 2.4: Agreement Labels[†]

Total disagreement

Total agreement

□ maximum　　×　average　　■ minimum
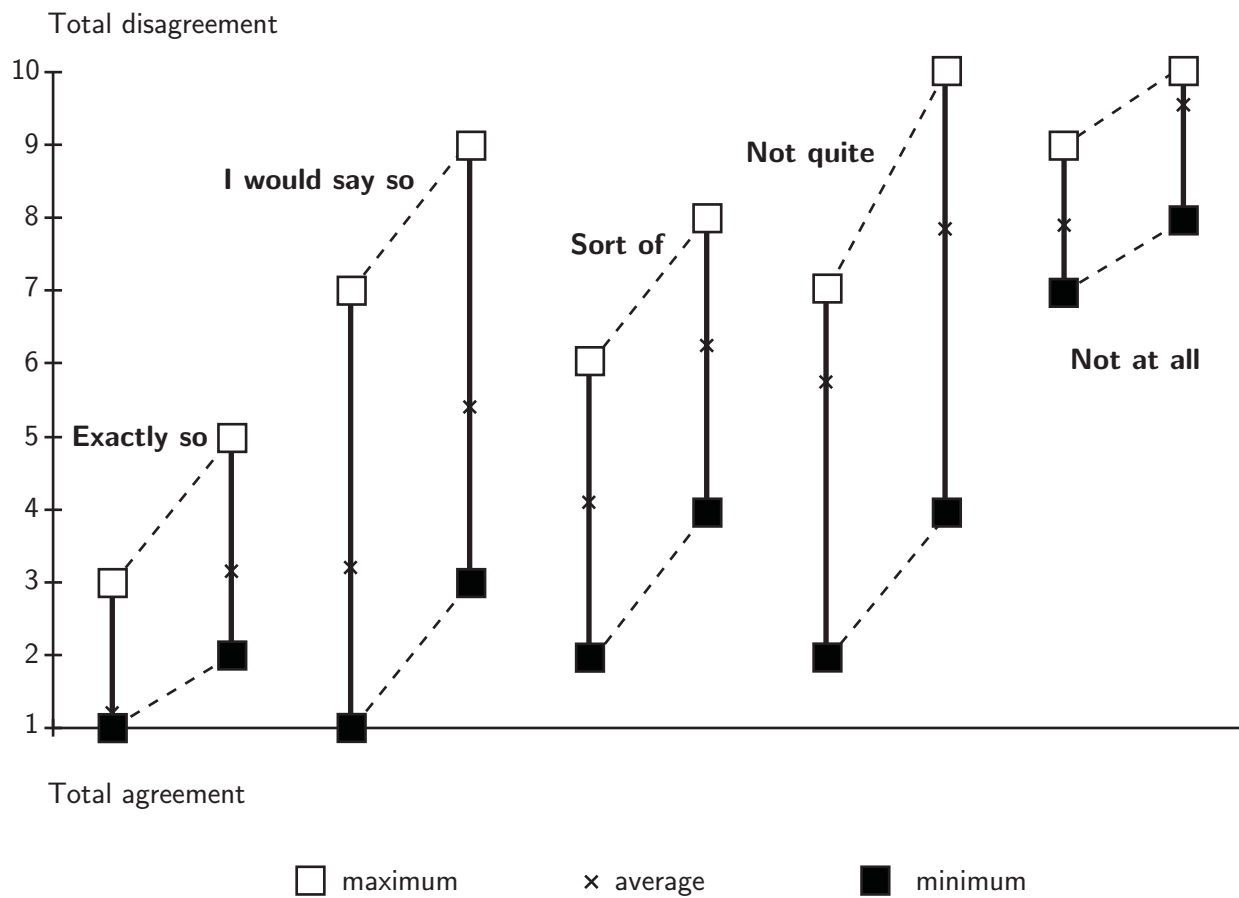
[†]For each wording label, the bar on the left is the lower limit, and the bar on the right is the upper limit of the label's position on the scale from 1 to 10.

precisely into a single category. That is, a discrete-choice model requires its choice options to be exactly defined and, accordingly, mutually exclusive. Consider the categories in the certainty follow-up question from Ready, Navrud, and Dubourg (2001):

I am *almost certain* (95%) that I would pay [that much money].

It is *more likely* that I *would* pay than that I would not.

It is *equally likely* that I would pay as that I would not.

It is *more likely* that I *would not* pay than that I would.

I am *almost certain* (95%) that I would *not* pay.

The meanings and boundaries of these categories are subject to personal interpretation. "More likely yes than no" seems to be broader a statement than "almost certainly yes" and may be thought of by some respondents as logically equivalent. The same applies to their negative counterparts. Finally, "equally likely" implies a point rather than a range.

Figure 2.4 illustrates this point. The author asked 20 students to place the intensity of five wordings for agreement or disagreement, as understood, on a scale from 1 to 10. In particular, the respondents were requested to indicate where in their mind each wording would begin and end, relative to the labels attached to the scale's end-points: total and unconditional agreement and radical disagreement. Wordings like these could have been response categories for a follow-up question: "Are you certain that you are willing to pay the stated amount?" As one can see from Figure 2.4, categories significantly overlap.

Measurement with verbal qualitative categories is problematic. First, verbal labels have no clear boundaries. An assumption that, for example, the label "small amount of money" covers any amount in the range from $1 to $25 would be as arbitrary as assuming the range from $5 to $50. Second, categories are subject to personal interpretation. As Mendel (2001) reiterates, words mean different things to different people. As a result, a particular feeling or impression cannot be put in a single category unambiguously and with certainty.

A remedy can be found in the theory of "fuzzy" sets. The key advantage is that fuzzy set theory allows partial membership of a point from a continuous set in many discrete

categories called fuzzy sets. This partial membership is represented through a membership function which can take any real value from zero to unity. Manipulation with fuzzy sets — the fuzzy logic — has found various applications in engineering; for example, in signal processing and automated control. Also, it can be and is used as an approximate reasoning tool in social research, especially when verbal discourse is involved (Zadeh 1975).

Van Kooten et al. (2001) devised a simple and elegant way to infer on WTP, using two fuzzy sets: a set labelled "Willingness to Pay" (WTP) and another labelled "Willingness Not to Pay" (WNTP). As the amount of the monetary bid grows, its membership in WTP decays and its membership in WNTP grows. An intersection of WTP and WNTP thus characterizes a reasonable locus of WTP values. Using the dataset from Li and Mattsson (1995), they reinterpreted the certainty follow-up question as a level of comfort for the individual with her dichotomous decision and, depending on whether it was "yes" or "no," as its membership in WTP or WNTP. Assuming a parametric relationship between the bid and membership values for both sets, Van Kooten et al. used a non-linear regression to estimate the parameters and then solved the system for the WTP value at which the membership in WTP is equal to that in WNTP.

In this dissertation, Case Study III (Chapter 5) presents a valuation methodology based on a model-free fuzzy logic advisor to perform mapping between inexactly defined categories of the Commodity Information Signal and Response Signals.

Respondents can attach different meanings to input and output categories. Further, respondents are almost certain to be heterogenous in their decision rules (Stored Rules Unit). The simplest and most commonly used way to account for respondents' decision-making heterogeneity is to include their observed demographic characteristics, such as gender, age, education, etc., in the valuation equation. This method, however, has two drawbacks. First, the investigator has no underlying theory as to the signs and magnitudes of the respective model parameters. Therefore, any significant estimate becomes acceptable. Second, if all commodity information were removed from the model, then the presence of respondent

characteristics would create two different utility levels for the same state of no commodity being consumed (see Case Study I).

With or without individual-specific factors, random utility models (RUM) have been used for a long time to account for the apparent variability of responses. Many economists and econometricians refer to a study by Thurstone (1927) as the origin of RUM. Thurstone described what he called the "law of comparative judgment." According to this law, an individual is volatile in his/her discriminal process, so that different comparative judgments are given on successive occasions about the same pair of items. Mathematically, the law was presented as

$$S_1 - S_2 = x_{12}\sqrt{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2} \tag{2.2}$$

where $S_j, j \in \{1,2\}$ are quantified attributes of the choice set elements (stimuli), σ_j^2 are their variances, r is the correlation coefficient, and a zero-centered x_{12} corresponds to the proportion of judgments $S_1 \succ S_2$.

All possible pairs constitute a "psychological continuum" and one of the stimuli serves as a point of reference. Despite the different terminology, one can clearly see the familiar parallels with utility. Importantly, Thurstone (ibid.) also introduced an assumption that apparent attributes of an item for a group of judges could be thought of as a random variable. Nevertheless, it is emphasized in the paper that one-time behavior of the group and repeated behavior of any its member cannot be equated.

The modern rendition of random utility translates the behavior of the group into that of a representative judge — the representative consumer. Hanemann (1984) states:

> A random utility model arises when one assumes that, although a consumer's utility function is deterministic for him, it contains some components which are unobservable to the econometric investigator and are treated by the investigator as random variables.

The modern approach to heterogeneity and randomness is different from Thurstone's study. There is one fundamental source of uncertainty in a modern RUM — the uncertainty is only related to unobservable factors that impact the respondent's decision-making. Every

respondent is assumed individually to exhibit a deterministic behavior. If the respondent herself becomes stochastic, then there are already two sources of uncertainty in the model, which cannot be separated by the investigator. While this fact has little implication when only one commodity is valued in an experiment, it may have fundamental consequences in understanding rationality of the respondent's choice when multiple items are sequentially valued in the same survey.

As an example, consider a T-period sequential binary choice model. A utility maximizer i chooses, at each period t, $t = 1 \ldots T$, in a sequence between two states of the world. These states are a period/individual-specific "alternative" (say, a higher level of an environmental good) and a baseline level of that good, the "status quo," with the corresponding utility levels:

$$u_{it} = v_{it} + \varepsilon_{it} \tag{2.3}$$

$$\tilde{u}_{it} = \tilde{\varepsilon}_{it}$$

where $v_{it} = v(\boldsymbol{x}_{it})$ is the deterministic utility of the alternative with attributes $\boldsymbol{x}_{it}$, the deterministic utility of the status quo is zero, and $(\varepsilon_{it}, \tilde{\varepsilon}_{it})$ are utility shocks.

Without imposing a special structure on the distribution of utility shocks, preferences of any given respondent cannot be considered completely reflexive or transitive. For example, $\tilde{\varepsilon}_{it}$ is a random variable. If it is continuous, then either $\tilde{\varepsilon}_{it} > \tilde{\varepsilon}_{is}$, $s \neq t$ or $\tilde{\varepsilon}_{it} < \tilde{\varepsilon}_{is}$; that is, the status quo is strongly preferred to itself for any pair of choices.

When multiple items are valued, a panel dataset arises. A majority of generic panel discrete-choice models ignore the above violations of utility axioms for the complete stochastic formulation of the model. Instead, only the deterministic utility v_{it} is assumed to adhere to utility theory.

Perhaps the most studied class of panel discrete-choice models is the error components specification. It is assumed that, apart from $(\boldsymbol{\varepsilon}_{it}, \tilde{\boldsymbol{\varepsilon}}_{it})$ that are independently and identically distributed variables, there is an additive term $\alpha_{it} \sim \mathrm{G}$ that reflects the idiosyncratic behavior of the individual given the choices. G is some multivariate distribution with zero mean.

Equation (2.3) thus becomes

$$
\begin{aligned}
u_{it} &= v_{it} + \alpha_{it} + \varepsilon_{it}, &&(2.4)\\
\tilde{u}_{it} &= \tilde{\varepsilon}_{it}.
\end{aligned}
$$

Two approaches to estimating such unobserved effects specification have long been standard in econometrics. One is a random effects (RE) approach. In its simplest form, it assumes that $\alpha_{i1} = \alpha_{i2} = \ldots = \alpha_{iT} = \alpha_i$ and $\alpha_i \sim \mathrm{G}(\boldsymbol{\theta}), \forall i$, where the univariate G is a specific, known distribution with unknown parameters. Then the probability of a particular sequence of outcomes is expressed as

$$
\Pr(1[u_{i1} \succ \tilde{u}_{i1}] = y_{i1}, 1[u_{i2} \succ \tilde{u}_{i2}] = y_{i2}, \ldots, 1[u_{iT} \succ \tilde{u}_{iT}] = y_{iT}) =
$$
$$
\mathrm{E}_\alpha \left[\prod_{t=1}^{T} \Lambda(v_{it})^{y_{it}} (1 - \Lambda(v_{it}))^{1-y_{it}} \right] \quad (2.5)
$$

where $\Lambda(\cdot)$ is a known link function and y_{it} is either 0 or 1. Although Equation (2.5) may have no closed-form expression on the right-hand side, estimation of the model is computationally straightforward with either Monte-Carlo simulated maximum likelihood, simulated moments, or with the use of Gaussian quadrature to perform numerical integration.

One may wish to relax the assumption that all effects for a given individual are equal and assume, instead, that $\boldsymbol{\alpha}_i \sim \mathrm{G}(\boldsymbol{\theta}), \forall i$, with the distribution of $\boldsymbol{\alpha}_i$ being now T-dimensional and of a known form, as before. Such models classify as a special case of random coefficient models (RCM).

RCM and its further generalization, such as random parameter models, allow immense flexibility (Walker 2001). By imposing a particular structure on parameters of the distribution of unobserved effects and their relationship to observed individual/choice characteristics, the researcher can tailor it to the peculiarities of the process being modelled. However, the distributional assumption itself is a very strong one. Whether or not unobserved effects actually follow the assumed distribution cannot be tested, nor is there usually any sound theory to justify the choice of a specific distribution. As a result, the estimates of parameters

in the model may have no useful properties at all, in case the *ad hoc* distributional assumption is wrong.

Another type of models that, too, can be considered error components models are fixed effects (FE) models[6]. The FE perspective views the additive unobserved effect in Equation (2.4) as an incidental parameter, whose population-wide behavior is both unknown and of no interest.

An example of an FE logit model/estimation technique was developed by Chamberlain (1980). For a model that features individual-specific heterogeneity (as in the simple RE case above), he suggested to maximize the conditional likelihood function

$$L^c = \prod_{i=1}^{N} \Pr\left(Y_{i1} = y_{i1}, Y_{i2} = y_{i2}, \ldots, Y_{iT} = y_{iT} \middle| \sum_{t=1}^{T} y_{it}\right) \tag{2.6}$$

where $Y_{it} = 1[u_{it} \succ \tilde{u}_{it}]$. The conditioning statistic is sufficient for the incidental parameter α_i. Chamberlain's method, however, does not use observations on those individuals for whom $\sum_{t=1}^{T} y_{it}$ is either 0 or T, allows only linear v_{it} and, naturally, excludes any choice attributes that remain constant for the individual.

Recent developments are often sophisticated hybrids of parametric and non-parametric models. A generalized estimating equation (GEE) model (Zeger et al. 1988) has recently received considerable attention. The GEE approach is to estimate the model from Equation (2.4) in the generalized linear model specification with a random effect:

$$E(y_{it}|\boldsymbol{x}_{it}) = \Lambda(\boldsymbol{x}_{it}\boldsymbol{\beta} + \alpha_{it}) \tag{2.7}$$

assuming that $\boldsymbol{\alpha}_i \sim G(\boldsymbol{0}, \mathbf{D})$. Although the distribution of the random effect is not directly specified, it is assumed to be of the linear exponential type. The estimation method itself is a complex form of generalized method of moments.

A unit non-response bias occurs when a survey respondent whose individual characteristics are available declines providing an answer to the elicitation question. Similarly, a

---

[6]In some textbooks, e.g. Train (2003), the term "error components" is only used in a random effects context.

respondent can refuse providing some of her demographic characteristics that are used in analysis. This problem is known as the item non-response.

There are two ways to deal with unit non-response or an incomplete response. One way is to remove incomplete observations from analysis and correct the model for a possible sample selection bias that may result in case the removal affects survey respondents in a systematic manner. Testing for a bias introduced by removing non-responses usually involves testing for no significant difference in respondents' characteristics in both parts of the sample (Messonnier et al. 2000). A variety of general econometric models with censoring or truncation can then be used to deal with the sample selection problem.

Consider, for example, a censored discrete-response model due to Wynand and van Praag (1981). In this model, a response to the elicitation question can only be observed if the respondent selects herself:

$$\Pr(outcome) = \begin{cases} \Pr(choice_j) \cdot p & \text{if } choice_j \text{ observed} \\ 1 - p & \text{if no response is given} \end{cases} \tag{2.8}$$

where $p = p(\mathbf{w}_i'\boldsymbol{\zeta})$, $\mathbf{w}_i$ is a vector of the respondent's personal characteristics, and $\boldsymbol{\zeta}$ are parameters. The above model of sample selection is implemented in Case Study I.

Another approach to tackling the problem of missing responses or individual attributes is through data imputation. Imputation refers to the replacement of missing data with a substitute that allows data analysis to be conducted. Typically, data imputation is conducted heuristically in two stages (Rubin 1987). First, the sample population is divided into several clusters where respondents have similar characteristics. Second, a donor unit is randomly chosen from the respective cluster and the replacement is carried out. Data imputation is, however, a controversial *ad hoc* technique, many of whose properties are unclear with heterogenous enough samples or depend on how the clustering was conducted.[7]

---

[7]An alternative approach to data imputation is through the use of the Expectation-Maximization (EM) algorithm (Bilmes 1998), which is a Markov chain Monte-Carlo technique. However, there have been very few successful implementations of the EM algorithm in economics to date.

Unlike the cases of heterogeneity and sample selection, there are no readily available econometric models to deal with warm glow/protest voting as well as many forms of framing. It does not appear to be technically feasible detecting these effects on an individual level without specially designed survey questions to identify problematic respondents. In case of warm glow, such a question may look like: "Do you think that every member of society should donate more for environmental restoration and protection [Yes/No]." A question similar to "Are you opposed to any increase in taxes, no matter what purpose the additional proceeds might be used for?" could be helpful in highlighting potential protest votes. Finally, re-phrasing or re-sequencing questions for a part of the sample can help with getting information on a particular kind of framing.

Once the additional explanatory variable or variables are available, these can be incorporated into the valuation model. Approaches and, consequently, results depend entirely on how researchers interpret the action of the effect. Some scholars consider it as depending on respondents/choice characteristics (Jorgensen et al. 1999). Nunes and Onofri (2004) provide an explanatory analysis of the individual warm glow motivational profile. They regress a warm glow motivation metric against a series of variables containing information about survey respondents characteristics.[8] The author would suggest a latent class model, which avoids using a warm glow/protest vote indicator in the valuation model directly or censoring suspected units completely, thus turning the case into a sample selection problem (Calia and Strazzera 1999).

Assume that each respondent is a warm-glower with probability p. A regular respondent answers "yes" to the valuation question with probability $\Pr(y_i = 1) = F(\boldsymbol{x}_i\boldsymbol{\beta})$; a warm-glower answers "yes" with probability 1 for any bid value. Assume also that each respondent is characterized by a binary, bivariate pair (z_i, w_i), where $z_i = 1$ indicates a warm-glower. The value of z_i cannot be observed for those respondents that answered "yes." A warm-glower answers "yes" to some warm-glow detection question with probability p_1; that is,

---

[8]Besides, Nunes and Onofri (2004) provide an extensive selection of "warm-glower" detection questions.

$\Pr(w_i = 1 | z_i = 1) = p_1$. A regular respondent answers "yes" to this question with probability p_2, such that $p_2 \leq p_1$. Now, the conditional likelihood of (y_i, w_i) pairs is given by $l_1^c = \left[w_i \left(p_1^{z_i} + (1 - p_1)^{1-z_i} \right) + (1 - w_i) \left(p_2^{z_i} + (1 - p_2)^{1-z_i} \right) \right] F(\boldsymbol{x}_i \boldsymbol{\beta})$ for all $y_i = 1$ and by $l_0 = \left[p_2^{z_i} + (1 - p_2)^{1-z_i} \right] (1 - F(\boldsymbol{x}_i \boldsymbol{\beta}))$ for all $y_i = 0$. This is a latent class model in z_i and its estimation is straightforward with the regular maximum likelihood by unconditioning the terms involving z: $l_1 = p l_1^c (z_i = 1) + (1 - p) l_1^c (z_i = 0)$. The model can be extended by assuming p to be a function of some attributes of the respondent's.

Herriges and Shogren (1996) model the anchoring effect in a bidding game experiment by representing WTP that the respondent uses at any bidding iteration, except the first one, as a linear combination of her true WTP and the bid amount from the previous iteration. As mentioned earlier, DeShazo (2002) uses prospect theory to model the anchoring effect. The author believes, however, that such models are very subjective and cannot be recommended for general application until enough empirical evidence amasses to corroborate their findings.

Another approach to dealing with warm glow/protest voting is to build more robust models or to use welfare change measures robust to a considerable number of outliers, which are warm-glowers and protest voters. A widely employed method is using the median of the estimated WTP distribution instead of its mean. This method advocated by Hanemann (1984) exploits the insensitivity of the median to abnormally high or low observations, as opposed to the distribution's estimated mean. Hanemann and Kanninen (1999) provide a short review of statistical literature on modelling parametric multi-modal distributions which could be useful for WTP distributions featuring warm glow "yeses" and protest "nays." An and Ayala (1996) suggest using a distribution mixture model.

2.4 Concluding Remarks

A number of substantial unintended effects have been described in this chapter and a practitioner-oriented review of approaches to modelling those effects has been conducted.

Some effects such as heterogeneity of respondents' decision-making rules or sample selection issues can be treated with standard econometric methods and models relatively easily. For other effect types that include warm glow, protest voting, and some kinds of framing, better survey design appears to be the preferred way to go, since an *a posteriori* modelling and treatment becomes highly subjective. Perception-related effects can be modelled with non-probabilistic devices.

An important note needs to be made at this point about rationality in the respondent's behavior. Theoretically, rationality is easily determined by whether or not a respondent adheres to the behavioral model the investigator has in mind. Empirically, the notion of rationality becomes blurry. Unintended effects do not occur spontaneously and inexplicably — they all happen for a reason. They happen because different people have individual and, sometimes, peculiar ways to make judgments. Everyone has one's own idiosyncratic logic and hence rationality. This creates a fundamental obstacle to modelling rationality.[9]

Different ways to model "rational behavior" are available. The first objective of the empirical Case Studies I–III in this dissertation is to demonstrate a broad range of possibilities for the applied environmental economist in the technology of stated-preference model building in the presence of unintended effects. Case Study I in Chapter 3 implements a form of RCM, a panel mixed logit model to account for respondents' heterogeneity in decision-making rules. The model also accounts for a sample selection bias and provides a treatment of "warm glow." Case Study II in Chapter 4 is generally dedicated to respondents' rationality in a sequential multiple commodity valuation setting. It starts with utility axioms and develops a detailed stochastic model which is consistent with the axioms of reflexivity, transitivity, and continuity. Finally, Case Study III in Chapter 5 presents a practical solution to the issue of impossibility of exactly quantifying qualitative categories in both commodity description and elicitation format options.

---

[9]At the very least, a doubt can be cast on the use of probabilistic reasoning in its modelling; see Moldoveanu and Langer (2002).

The second objective of the Case Studies is to illustrate model implementation using nonstandard mathematical methods. Case Study I employs an evolutionary algorithm, which can be used in optimization with complex econometric models. Case Study III presents a non-probabilistic valuation technique based on a model-free fuzzy expert system for social judgments.

CHAPTER 3

CASE STUDY I

VALUATION OF FARMLAND CONSERVATION EASEMENT PROGRAMS WHEN

PREFERENCES VARY[1]

---

[1]Volinskiy, D., and J. Bergstrom. To be submitted to *Land Economics.*
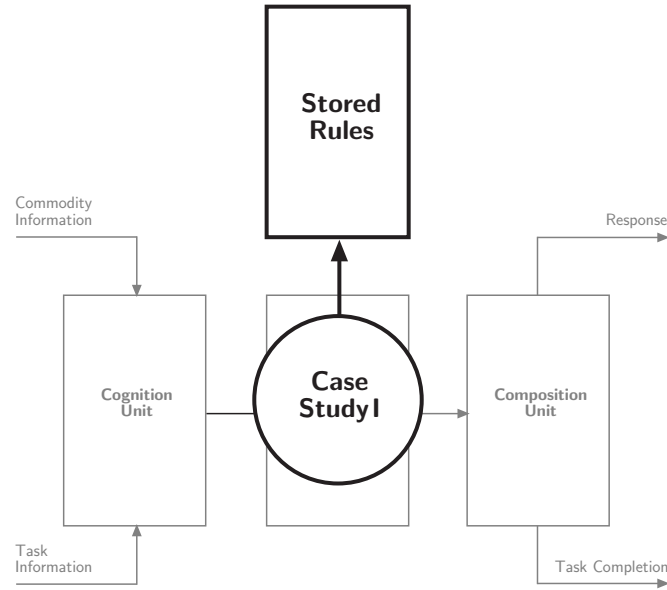
3.1 ABSTRACT

*Case Study I looks at two types of unintended effects that occur in the **Stored Rules** unit of the respondent's decision-making system. These are effects of respondents heterogeneity on parameters of rules and partial non-responses. People are different while environmental commodities are complex. Not only will there be a cumulative effect of the respondent's unobserved idiosyncracies on choices, major economic parameters of interest, such as rates of substitution and money-metric marginal utilities are likely to vary from person to person. As an illustration of how to deal with the variation in tastes and choices, Case Study I develops and implements a random parameters model to estimate people's willingness-to-pay for farmland conservation easement programs. The model allows for conservation preferences and tastes to vary randomly over people. The model also makes a correction for sample selection bias by allowing preferences to answer valuation questions to be random. Finally, the model accounts for an effect of "warm glow."*

3.2 Introduction and Objectives

Marked with a rapid pace of urban sprawl, the last decades saw a broad public concern that vast expanses of prime farmland could be irrevocably lost due to commercial and residential conversion. Georgia's 40,000 farms cover nearly a third of the state's land area. As the growth of Atlanta and other city conglomerates continues, sprawling development is threatening to decimate Georgia's agricultural lands, which have traditionally been at the core of this Southern state's economy. Voters in the Southeast are amongst the most concerned in the nation about disappearing farmland — 70% worry that too much farmland is being destroyed by development (Esseks 2001). According to American Farmland Trust (Sorensen, Greene, and Russ 1997), Georgia ranked third nationally during 1992–1997 for productive agricultural acreage lost to paved surfaces.

Following the enactment of the Farmland Protection Policy Act of 1981 and subsequent regulations, USDA instituted a Farm and Ranch Land Protection Program (FRPP). The program provides funds to help purchase development rights in order to keep productive farmland in agricultural use. Purchasing an easement compensates the land owner for forfeiting the right to use the land for any purpose. The farmer's rights to allow or conduct non-agricultural development of the land become permanently limited, while all other property rights pertaining to that land remain intact. The money available to the farmer through the easement contract can partially cover the opportunity costs of using the land for agriculture, creating a counter-balance to offers from commercial land markets.

A conservation easement program is essentially a Coasean solution where the government facilitates the re-assignment of property rights. The prime objective of the policy-maker is to parameterize the solution in such a way that tax payers' welfare is maximized. The policy-maker should determine the size of the program, characteristics of the farmland to target, and how much people would be willing to pay to provide the government with the money to purchase easements with that program. There is no private market for development rights on

farmland. Consequently, their valuation is often conducted using stated-preference surveys of the general population.

Recent stated-preference studies of farmland preservation benefits (Bowker and Didychuk (1994), Kline and Wichelns (1998), Lavigno (2004), Paterson et al. (2004)) assumed that different survey respondents with the same observed characteristics had the same value for the preserved farmland quantity and characteristics. The assumption of identical tastes does not seem to be very realistic given the multitude of services that natural and manmade agricultural environments provide (Bergstrom, Dillman, and Stoll 1985). This study develops and implements a random parameters model to estimate Georgians' willingness-to-pay (WTP) for farmland conservation easement programs, which allows for conservation preferences and tastes to vary randomly over people. The model also makes a correction for sample selection bias by allowing preferences to answer valuation questions to be random as well. Finally, the model accounts for an effect of "warm glow."

In what follows, Section 3.3 introduces the theoretical background of multi-attribute analysis of programs, Section 3.4 presents a random coefficient stochastic model, Section 3.5 introduces the study data and discusses the selection of variables, Section 3.6 contains technical details of the optimization process, Section 3.7 comments on estimation results, and Section 3.8 concludes with recommendations on the optimum conservation easement programs and outlines further improvements to be made to the model.

3.3 Multi-Attribute Contingent Valuation Model

We use what we term a multi-attribute contingent valuation method (MCVM) for this study. Often used under different names, MCVM occupies a middle ground between the closed-ended format of the contingent valuation method (CVM) and conjoint analysis techniques (Boxall et al. 1996) by varying commodity attributes along with the price and quantity in a multiple discrete-choice valuation format. In an MCVM application, respondents are asked to indicate their preferred choice out of two or more alternative states of the world,

further referred to as alternatives, that have differing qualitative and quantitative attributes, as well as costs.

A respondent is assumed to choose Alternative 1 over Alternative 0 if and only if

$$v_1(\mathbf{q}_1, I - P, \mathbf{s}) > v_0(\mathbf{q}_0, I, \mathbf{s}) \tag{3.1}$$

where $v_1(\cdot), v_0(\cdot)$ are her utility levels attainable with (vector-valued) alternative characteristics, $\mathbf{q}_1$ and $\mathbf{q}_0$, $\mathbf{q}_1$'s net price of P, and constant income I and other factors $\mathbf{s}$.

If the indirect utility function is separable and linear, $V(\mathbf{q}_j, I_j) = \langle \mathbf{z}_j | \boldsymbol{\alpha} \rangle + \gamma I_j, j \in \{0, 1\}$, then it is straightforward to show, by equating the utility levels and solving for price, that the compensating surplus measure for changing from Alternative 0 to Alternative 1 is

$$P^* = CS = \frac{\langle \boldsymbol{\alpha} | (\mathbf{z}_1 - \mathbf{z}_0) \rangle}{\gamma} \tag{3.2}$$

where $\langle \mathbf{z}_1 | \boldsymbol{\alpha} \rangle$ and $\langle \mathbf{z}_0 | \boldsymbol{\alpha} \rangle$ are linear functions of attributes z_1 and z_0 of Alternatives 1 and 0. Notably, Equation (3.2) represents the optimum consumption condition, under which the marginal rate of substitution is equal to the price ratio.

Put into a random utility context, Equation (3.1) becomes

$$\begin{aligned} \Pr(v_1 > v_0) &= \Pr(\langle \boldsymbol{\alpha} | (\mathbf{z}_1 - \mathbf{z}_0) \rangle - \gamma P > \epsilon_0 - \epsilon_1) \\ &= F_{\epsilon_0 - \epsilon_1}(\langle \boldsymbol{\alpha} | (\mathbf{z}_1 - \mathbf{z}_0) \rangle - \gamma P) \end{aligned} \tag{3.3}$$

where $v_j = \langle \mathbf{z}_j | \boldsymbol{\alpha} \rangle + \gamma I_j + \epsilon_j, j \in \{0, 1\}$, ϵ_j is a disturbance component, and $F_{\epsilon_0 - \epsilon_1}(\cdot)$ is the distribution function of the difference of disturbances. Upon making a parametric assumption about $F_{\epsilon_0 - \epsilon_1}(\cdot)$, parameters $\boldsymbol{\alpha}$ and γ of the model can be estimated and the welfare measure computed according to Equation (3.2) (Hanemann 1984).

3.4 Empirical Methodology

The workhorse of multiple discrete-choice models is the multinomial logit model (MNL). MNL choice probabilities are obtained as:

$$\Pr(Y_i = j) = \frac{\exp(\mathbf{x}_j' \boldsymbol{\beta})}{\sum_{J_i} \exp(\mathbf{x}_j' \boldsymbol{\beta})} \tag{3.4}$$

where Y_i is the index of the alternative chosen by an individual i out of a choice set that contained J_i alternatives; $\mathbf{x}'_j = [\mathbf{z}_j \; P]$ are vectors of attributes of alternatives in that choice set, and $\boldsymbol{\beta}' = [\boldsymbol{\alpha} \; \gamma]$ is a vector of the individual's characteristics.

MNL is often used with data from multiple discrete-choice valuation surveys. One of the key assumptions of the model's is that individuals both in the sample and population-wide have identical characteristics; that is, $\boldsymbol{\beta}_i = \boldsymbol{\beta} \; \forall i \in \Omega$. This assumption effectively allows one to pool all available observations and estimate Equation (3.4) by the regular maximum likelihood. However, this assumption is hardly a realistic one. Respondents come from different cultural backgrounds, have different social and political agenda. Such important parameters as substitution patterns and money-metric utilities of alternatives' attributes are unlikely to be the same for everybody. With the parameters of interest being random, a powerful modelling framework is offered by what is known as mixed models.

A mixed model relaxes the assumption of respondents being identical, replacing it with a much milder one of respondents being identically distributed. That means $\boldsymbol{\beta}_i \sim \mathrm{D}(\boldsymbol{\theta}) \; \forall i \in \Omega$, where $\mathrm{D}(\cdot)$ is some multivariate distribution of a known form.

If the distribution of the shock is a generalized extreme value (GEV) distribution, the model is called a mixed logit model. Applications of mixed logit in environmental and resource economics appeared with an advent of accessible high-performance computing. The most common form of mixed logit is a random coefficient MNL (RCMNL), which is the case when the decision-maker's objective function is linear in parameters and shocks are independently and identically distributed. McFadden and Train (2000) argue that an RCMNL model can approximate any reasonable utility-maximizing behavior. Train (1999) used panel RCMNL to model demand for angling sites. Siikamaki (2001) applied a similar model to value bio-diversity preservation programs. Train (2003) provides additional discussion of RCMNL applications in other fields of economics.

The probability of making a particular choice is expressed in RCMNL as an expectation over $\boldsymbol{\beta}$ of the right-hand side of Equation (3.4):

$$\Pr(Y_i = j) = \mathrm{E}_{\boldsymbol{\beta}}\left[\frac{\exp(\mathbf{x}_j'\boldsymbol{\beta})}{\sum_{J_i} \exp(\mathbf{x}_j'\boldsymbol{\beta})}\right] \tag{3.5}$$

The RCMNL framework can be easily extended to accommodate panel data. Assuming that responses are longitudinally uncorrelated and given the values of the individual's coefficients, the probability that the respondent i makes a sequence of choices over the periods T_i is given by

$$\ell_i(\boldsymbol{\theta}, \mathbf{x_i}) = \mathrm{E}_{\boldsymbol{\beta}}\left[\prod_{t=1}^{T_i} \frac{\exp(\mathbf{x}_{jt}'\boldsymbol{\beta})}{\sum_{J_i(t)} \exp(\mathbf{x}_{jt}'\boldsymbol{\beta})}\right] = \mathrm{E}_{\boldsymbol{\beta}}\left[\prod_{t=1}^{T_i} \Lambda_{\mathrm{MNL}}(\mathbf{x_{it}}, J_i(t)|\boldsymbol{\beta})\right] \tag{3.6}$$

where $J_i(t)$ indexes the set of choices available to the respondent at time t, and $\Lambda_{\mathrm{MNL}}(\cdot)$ is used to denote MNL choice probabilities (Equation (3.4)), to simplify notation.

Stated preferences are revealed preferences of the agent's hypothetical "self." People have different attitudes and abilities with respect to planning. This implies that the degree to which the agent can associate her actual "self" with a hypothetical one is an unknown variable factor. If it is related to a subset of parameters in the conditional utility, a selection problem arises. RCMNL allows to correct for a possible selection problem in a simple and logical way.

Suppose that the degree of association between the respondent's "selves" can be approximated by a parametric function of some of the respondent's observed characteristics and an unobserved shock. Also suppose that a response can only be observed only if an individual-specific threshold level is reached. This is a censored discrete response model (Wynand and van Praag 1981; Greene 1992):

$$\Pr(outcome) = \begin{cases} \Pr(choice_j) \cdot p & \text{if } choice_j \text{ observed} \\ 1 - p & \text{if no response is given} \end{cases} \tag{3.7}$$

where $p = p(\mathbf{w}_i'\boldsymbol{\zeta})$, $\mathbf{w}_i$ is a vector of the respondent's personal characteristics, and $\boldsymbol{\zeta}$ are parameters.

Let $\iota(Y_{it})$ be an indicator of an event that Y_{it} is observed. Let us assume that the probability of selection follows a standard binary logit model; that is, $\Pr(\iota(Y_{it}) = 1) = \Lambda(\mathbf{w}_i'\boldsymbol{\zeta}_i)$, where $\Lambda(\cdot)$ is the binary logistic distribution function, and let $[\boldsymbol{\beta}_i \; \boldsymbol{\zeta}_i]' \sim \mathrm{D}(\boldsymbol{\theta}) \; \forall i \in \Omega$. With these modifications Equation (3.6) can be written as

$$
\ell_i(\boldsymbol{\theta}, \mathbf{x_i}, \mathbf{w}_i) = \mathrm{E}_{\boldsymbol{\beta},\boldsymbol{\zeta}} \left[\prod_{t=1}^{T_i} \left(\left\{ \Lambda_{\mathrm{MNL}}(\mathbf{x_{it}}, J_i(t)|\boldsymbol{\beta})\Lambda(\mathbf{w}_i'\boldsymbol{\zeta}) \right\}^{\iota(Y_{it})} \right. \right.
$$
$$
\left. \left. + \left\{ 1 - \Lambda(\mathbf{w}_i'\boldsymbol{\zeta}) \right\}^{1-\iota(Y_{it})} \right) \right] \tag{3.8}
$$

A stated-preference model needs to account for a possible "warm glow" effect. This effect arises when the person derives additional utility from contributing to the provision of a public good (Andreoni 1990). If exhibited, "warm glow" will increase the probability of voting "yes" and thus bias WTP estimates upwards. "Warm glow" can also be reversed. For example, a respondent may vote against because she does not want to concede any new taxes in the locality or because of the opposition to liberal environmentalism. Either is a case of protest voting. One can see that "warm glow" and some cases of protest voting similarly relate to a political or ethical stance effect in one's preferences.

If a generalized "warm glow" effect wg is additive, then either

$$
v_i = wg + \langle \boldsymbol{\alpha}|\mathbf{z}_i \rangle - \gamma P \tag{3.9a}
$$
$$
v_0 = 0
$$

or

$$
v_i = \langle \boldsymbol{\alpha}|\mathbf{z}_i \rangle - \gamma P \tag{3.9b}
$$
$$
v_0 = wg
$$

This means that the "warm glow" can be modelled as either constant in the conditional utility for all alternatives, with that of the baseline being normalized to zero, or as an attribute of the baseline with no constant put in the conditional utility of any alternative.

With a parametric assumption that $[\boldsymbol{\beta}\ \boldsymbol{\zeta}]' \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we obtain the complete specification of the expected likelihood in Equation (3.8):

$$\ell_i = \iint \cdots \int \prod_{t=1}^{T_i} \bigg(\big\{ \Lambda_{\mathrm{MNL}}(\mathbf{x_{i}}_t, J_i(t)|\boldsymbol{\beta}) \Lambda(\mathbf{w}_i'\boldsymbol{\zeta}) \big\}^{\iota(Y_{it})}$$
$$+ \big\{ 1 - \Lambda(\mathbf{w}_i'\boldsymbol{\zeta}) \big\}^{1-\iota(Y_{it})} \bigg) \phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\boldsymbol{\beta}, \boldsymbol{\zeta}) \mathrm{d}\boldsymbol{\beta}\mathrm{d}\boldsymbol{\zeta}$$
$$= \iint \cdots \int (\ell_i|\boldsymbol{\beta}, \boldsymbol{\zeta}) \phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\boldsymbol{\beta}, \boldsymbol{\zeta}) \mathrm{d}\boldsymbol{\beta}\mathrm{d}\boldsymbol{\zeta} \tag{3.10}$$

where $\phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\cdot)$ is the multivariate normal density of $[\boldsymbol{\beta}\ \boldsymbol{\zeta}]'$ with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

Since it is not feasible to perform the multifold integration analytically, methods available to estimate RCMNL models — simulated maximum likelihood (SML), simulated moments (MSM) and scores (MSS) — are simulation-assisted. These methods can in theory handle any dimensions. The traditional Gaussian quadrature (GQ) is often deprecated as limited to lower dimensions (Evans and Swartz 1995). However, most practical problems are those involving comparatively low dimensions. GQ in such cases is quite fast and precise. Moreover, GQ is meaningful because it links RCMNL and latent class models (Greene 2001).

GQ would approximate ℓ_i by discretizing the multivariate normal distribution; that is

$$\tilde{\ell}_i = \sum_{r=1}^{R} \omega^r (\ell_i|\{\boldsymbol{\beta}, \boldsymbol{\zeta}\}^r) \tag{3.11}$$

where nodes $\{\boldsymbol{\beta}, \boldsymbol{\zeta}\}^r$ and weights ω^r are computed so as to match $R^{\frac{1}{d}}$ raw moments of the d-dimensional $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

The simulated or approximated log-likelihood function

$$\ln\tilde{\mathcal{L}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}|\mathbf{X}, \mathbf{W}) = \sum_{i=1}^{N} \ln\tilde{\ell}_i \tag{3.12}$$

where N is the sample size, can be maximized with any appropriate optimization method. An experiment on generated data showed that, *ceteris paribus*, GQ was roughly 2.5 times slower than SML, but it provided a gain in precision of 12 orders of magnitude. The optimization algorithm outlined in Section 3.6 makes use of GQ approximation.

3.5 Model Data and Variable Selection

The survey data for this study were made readily available courtesy of a project funded through USDA National Research Initiative[2], that aimed to identify the most important attributes of the farmland in Colorado, Georgia, Maine, Ohio, and Oregon to target in conservation easement programs. Details of the survey organization and composition can be found in Ozdemir (2003, Appendix B).

The dataset for Georgia includes responses from 213 individuals. Each individual in the survey was presented with four pairs of conjoint questions. Each question pair consisted of choosing "Alternative A vs. Alternative B" and "Alternative A vs. Alternative B vs. Neither Alternative," the latter corresponding to v_0 in Section 3.3. Attributes of alternatives are presented in Table 3.1. The survey also included some demographic information. In particular, the respondent's age, gender, years in school, number of people in the household, and income range were recorded.

If coefficients on all alternative attributes are considered to be random, the model may become inestimable. Estimation of a k-dimensional distribution entails optimization with respect to $k(3k+1)/2$ parameters, which is a computational task of overwhelming size when k grows large. RCM is easily fit for the estimation of constant coefficients as well — one may think of those as coming from a degenerate distribution. But choosing which coefficients are likely to be random and which constant represents a considerable problem.

Unfortunately, the only way to choose is by doing that *ad hoc*. Attributes COST and ACRES, i.e. the price and quantity, are central to shaping the individual's demand and thus likely to be quite diverse in the population. Besides, these two constitute the necessary minimum of program attributes and therefore determine the functional form of WTP distribution

Table 3.1: Attributes of Choice Alternatives

Attribute Mnemonic	Coding and Description
COST	One-time cost to household, US$; takes values in $-\{0.03, 0.05, 0.07, 0.1, 0.25, 0.5\} \times 10^2$
ACRES	Total acres of easements to purchase; takes values in $\{0.1, 0.5, 1, 2\} \times 10^6$
LOC	Farmland location priority; 1 if near urban areas, 0 otherwise
QUAL	Land quality priority; 1 if prime farmland, 0 otherwise
GRAIN	1 if land used for crops, 0 otherwise
HAY	1 if land used for hay, 0 otherwise
PASTURE	1 if land used as pasture, 0 otherwise
VEGET	1 if land used for vegetables, fruit, nuts, etc. 0 otherwise
FOREST	1 if land used for timber, 0 otherwise

for a program with minimal attributes. Accordingly, coefficients on COST and ACRES were to be included in the elicitation equation as random parameters.

In order to avoid negative WTP values, the distribution of the coefficient on COST is typically assumed to be log-normal (Train 1999), mainly to simplify computations[3]. However, it was decided not to force the coefficient to be positive on the premise that normal distribution provides sufficient flexibility to have this coefficient predominantly positive, if the data support it. The distribution of the coefficient on ACRES was not transformed, either.

Another random coefficient was introduced to capture the "warm glow" effect (Barron and Greene 1996). Finally, a random coefficient was used to reflect a random effect in the selec-

---

[3]Researchers interested in WTP distributions, e.g. McFadden (1994), An and Ayala (1996), have investigated a variety of functional forms, but those distributions came from the respective parametric assumptions about WTP, not from distributions of marginal prices.

Table 3.2: Model Parameters

Coefficient on:	Description
Random Coefficients, $\boldsymbol{\beta}$	
COST	One-time cost to household, US$
ACRES	Total acres of easements to purchase, million
WG	"Warm glow" random parameter
Random Coefficients, $\boldsymbol{\zeta}$	
RESEL	Random effect in selection equation
Constant Coefficients, $\underline{\boldsymbol{\beta}}$	
HFOOD	Human food production use
AFOOD	Animal food production use
FOREST	Timber production use
LOC	Farmland location priority
QUAL	Land quality priority
Constant Coefficients, $\underline{\boldsymbol{\zeta}}$	
SEX	Respondent's gender
AGE	Respondent's age
COLLEGE	Indicator of college education

tion equation, i.e. to capture unobserved personal characteristics that might have affected the selection of responding individuals. The model effectively became random effects logit. All farmland attributes except for the price and program size were considered constant coefficients, not subject to a considerable taste variation.

As part of the survey, respondents were asked to indicate the importance of land use attributes, GRAIN–FOREST, on a scale from one to seven. These weights were used to test whether respondents differentiated amongst all types of land use. Friedman's test, also known as the nonparametric ANOVA (Conover 1999), was conducted. The test returned

an F-distributed statistic of 9.7, significant at 90% confidence level, which indicated that different types of land use should have had some effect on responses. Post-test comparisons by Dunn's technique revealed no significant difference between the following groups of attributes: {GRAIN, VEGET}, {HAY, PASTURE}, and FOREST. This grouping has a lot of intuitive appeal, for the first pair belongs to human food production category, the second is animal food, and the last one is neither. Correspondingly, land use attributes entered the model grouped grouped into the three categories.

Personal attributes: age (AGE), gender (SEX), and an indicator of whether the respondent had college education (COLLEGE) were included in the selection equation as constant coefficients. The income range was not included due to the very imprecise nature of this factor. The estimation of the covariance matrix $\boldsymbol{\Sigma}$ of $\begin{bmatrix} \boldsymbol{\beta} & \boldsymbol{\zeta} \end{bmatrix}'$ was set up through its Cholesky factor $\mathbf{U}$ ($\mathbf{U}'\mathbf{U} = \boldsymbol{\Sigma}$). Optimization was to produce estimates of 14 ($k = 4$) parameters of $\mathrm{N}(\boldsymbol{\mu}, \mathbf{U})$ distribution of $\begin{bmatrix} \boldsymbol{\beta} & \boldsymbol{\zeta} \end{bmatrix}'$ and also estimates of 8 constant coefficients in $\begin{bmatrix} \underline{\boldsymbol{\beta}} & \underline{\boldsymbol{\zeta}} \end{bmatrix}'$, which, altogether, added up to 22 free parameters to estimate. Table 3.2 summarizes the model parameters.

3.6 OPTIMIZATION

Even with the number of random coefficients reduced, optimization of the log-likelihood function from Equations (3.10)–(3.12) represents a challenging computational exercise, given 22 parameters to estimate. Apart from the dimensionality issue, another problem is that every single call of the log-likelihood function is very computationally expensive, since it involves numerical integration. As a result, common quasi-Newton and simplex algorithms were found to perform quite poorly, ultimately grinding to a halt.

As an attempt to get around these obstacles, an evolutionary optimization algorithm was devised and implemented. Evolutionary algorithms offer a number of significant advantages over algebraically- or geometrically-derived ones: they are not susceptible to converging at local extrema because of a bad starting point nor do they require any smoothness or regularity

of the objective function; complexity does not grow dramatically with increasing dimensions, and numerical problems are less of a pitfall. The downside, however, are the absence of clear convergence criteria and the necessity to generate a lot of random events.

An evolutionary (stochastic search) algorithm takes after the concept of biological evolution and includes the following components (Eiben and Smith 2003):

(a) Create a generation of parents: candidate optimal values of the optimization parameter;

(b) Perform crossover by selecting mating pairs;

(c) Obtain offspring as a function of parameters of parents;

(d) Let the offspring mutate by perturbing their parameter values by random quantities;

(e) Rank the offspring by their fitness values, i.e. their objective function values;

(f) Combine the generations and implement survival of the fittest by shedding the least fit individuals; Repeat the cycle until an offspring generation becomes homogeneous enough according to a chosen criterion.

An informal description of an algorithm implemented for this study follows.

Step 0: Initialize by setting the counter $c = 0$, selecting the maximum number of generations $genMax$, the number of individuals of each sex, n, the number of immigrants of each sex, $\ddot{n}$, mutation step size $\kappa^{(c)}$, its lower and upper limits, κ_{min} and κ_{max}, mutation parameter $p_\kappa \in [0.817, 1.0)$, and choosing tolerance, τ, to be some small number.

Step 1: Create an initial (parent) generation of candidate parameter vectors, n males and n females, by drawing $\{\mathbf{B}_m^{(c)}, \mathbf{B}_f^{(c)}\}$, where each $\mathbf{B}$ has 22 rows and n columns, from the i.i.d. standard normal distribution.

Step 2: Assess fitness of the parents by evaluating the log-likelihood function for each column.

Step 3: Rank parents highest to lowest, within each sex, by rearranging the respective columns of $\mathbf{B}$'s.

Step 4: If $\kappa^{(c-1)} < \kappa_{min}$, reshuffle parents by randomly swapping columns of $\mathbf{B}_m^{(c)}, \mathbf{B}_f^{(c)}$ within each sex several times.

Step 5: Obtain the top male, $T_m^{(c)}$, and the top female, $T_f^{(c)}$, which are the individuals with the highest fitness values in their respective sex groups. Check the "soft" convergence criterion, $\|T_m^{(c)} - T_f^{(c)}\| < \tau$, and the "hard" one, $c > genMax$. Exit if either is satisfied.

Step 6: Create offspring, n males and n females, through a weighted average of parents (heuristic crossover): $\dot{\mathbf{B}}_{m_i}^{(c)} = \mathbf{U}_{m_i}^{(c)} * \mathbf{B}_{m_i}^{(c)} + (\mathbf{1} - \mathbf{U}_{m_i}^{(c)}) * \mathbf{B}_{f_i}^{(c)}$, $\dot{\mathbf{B}}_{f_i}^{(c)} = \mathbf{U}_{f_i}^{(c)} * \mathbf{B}_{m_i}^{(c)} + (\mathbf{1} - \mathbf{U}_{f_i}^{(c)}) * \mathbf{B}_{f_i}^{(c)}$, where $*$ denotes Hadamard product, $\mathbf{U}_{m_i}^{(c)}, \mathbf{U}_{f_i}^{(c)}$ are draws from the i.i.d. uniform on $[a, b]$, such that $a = 0.5, b = 1.0$ if $\mathbf{B}_{m_i}^{(c)} > \mathbf{B}_{f_i}^{(c)}$ and $a = 0.0, b = 0.5$ if $\mathbf{B}_{m_i}^{(c)} < \mathbf{B}_{f_i}^{(c)}$, $\forall i = 1, 2 \ldots n$ columns.

Step 7: Mutate the offspring by adding a multiple of a draw from the standard normal to offspring individuals, $\dot{\mathbf{B}}_s^{(c)} := \dot{\mathbf{B}}_s^{(c)} + \kappa^{(c)} \cdot \mathbf{M}_s + \mathbf{\Psi}_s^{(c)}$, $s \in \{m, f\}$, where $\mathbf{M}_s$ is a draw from the i.i.d. standard normal, and $\mathbf{\Psi}_s^{(c)}$ is a shift from Step 12.

Step 8: Assess fitness of the offspring and calculate, p_{Rech}, the percentage of successful individuals, i.e. those whose fitness exceeds the fitness of the parents.

Step 9: Adjust mutation step size according to the "1/5 success rule" (Rechenberg 1973). If $p_{Rech} < 20\%$, then $\kappa^{(c+1)} = \kappa^{(c)} \cdot p_\kappa$, and, if $p_{Rech} > 20\%$, then $\kappa^{(c+1)} = \kappa^{(c)}/p_\kappa$. Check for bounds: if $\kappa^{(c+1)} > \kappa_{max}$, set $\kappa^{(c+1)} = \kappa_{max}$, and, if $\kappa^{(c+1)} < \kappa_{min}$, set $\kappa^{(c+1)} = \kappa_{min}$ and reshuffle parents in $c + 1$-th generation (Step 3).

Step 10: Merge the parents and the offspring, $\mathbf{B}_s^{(c)} := [\mathbf{B}_s^{(c)} \; \dot{\mathbf{B}}_s^{(c)}]$, and rank each sex by fitness, as above.

Step 11: Select most fit $n - \ddot{n}$ individuals from each sex, and store them in $\mathbf{B}_s^{(c)}$, respectively; discard the rest.

Step 12: Calculate shift, $\boldsymbol{\Psi}_s^{(c+1)}$: if, in $\mathbf{B}_s^{(c)}$, some candidate parameter k monotonically increases/decreases, set the k-th row of the respective $\boldsymbol{\Psi}_s^{(c+1)}$ to $\frac{1}{n-\ddot{n}-1}$ of its range.

Step 13: Generate i.i.d. standard normal immigrants, $\ddot{\mathbf{B}}_s^{(c)}$, one for each sex, obtain their fitness values, and append the immigrants, $\mathbf{B}_s^{(c)} := \left[\mathbf{B}_s^{(c)}\ \ddot{\mathbf{B}}_s^{(c)} \right]$.

Step 14: Complete the new generation, $\mathbf{B}_s^{(c+1)} = \mathbf{B}_s^{(c)}$, increment the counter, $c := c+1$, and go to Step 3.

Information from 195 respondents was available to calculate the objective log-likelihood; 625 (5 in each dimension) evaluation nodes were used in GQ. The algorithm was initialized with: $genMax = 10000, n = 5, \ddot{n} = 1, \kappa^{(1)} = 0.25, \kappa_{min} = 0.01, \kappa_{max} = 1.0, p_\kappa = 0.817, \tau = 10^{-3}$. The optimization was completed in a total of 1142 generations; the value of the simulated log-likelihood function at the optimum was -1150.

3.7 Estimation Results and Discussion

The estimated means and standard deviations (obtained from $\hat{\boldsymbol{\Sigma}} = \hat{\mathbf{U}}'\hat{\mathbf{U}}$) of the random coefficients distribution, as well as estimated constant coefficients are presented in Table 3.3. Table 3.4 contains an estimate of the correlation matrix.

Examining the estimates, one can see that the panel was quite heterogeneous, hence a large variation in random coefficients. The coefficient on cost is mainly positive in the population; that on acres is centered near zero, which implies a comparatively low effect of program size, alone. However, with the favorable attributes: human food production use, prime farmland near urban areas, the numerator in the WTP equation becomes predominantly positive, especially for smaller program sizes.

Land use priorities for human and animal food uses have almost the same weights but in different directions; targeting the timber use is favored a little bit less. This finding is roughly consistent with a WTP aggregation in Feather and Barnard (2003), which places almost equal priorities on cropland and forest land and a smaller (but still positive) weight

Table 3.3: Estimated Coefficients

Coefficient on:	Mean (std.dev)	Standard deviation (std.dev)
Random Coefficients, $\boldsymbol{\beta}$		
COST	7.1152***	7.6988***
	(0.7835)	(0.8360)
ACRES	0.4953***	1.6907***
	(0.1313)	(0.1715)
WG	−4.2210***	8.3275***
	(1.0648)	(1.9252)
Random Coefficients, $\boldsymbol{\zeta}$		
RESEL	10.0000***	5.0584***
	(3.0272)	(0.6173)
Constant Coefficients, $\underline{\boldsymbol{\beta}}$		
HFOOD	0.4103***	—
	(0.1462)	
AFOOD	−0.4208***	—
	(0.1391)	
FOREST	0.3667**	—
	(0.1559)	
LOC	0.4115***	—
	(0.0846)	
QUAL	0.7385***	—
	(0.0801)	
Constant Coefficients, $\underline{\boldsymbol{\zeta}}$		
SEX	0.3966	—
	(0.5632)	
AGE	−0.5549	—
	(2.7463)	
COLLEGE	0.1468	—
	(1.0944)	

***—significant at 99% level; **—significant at 95% level.

Table 3.4: Correlation of Random Coefficients

	COST	ACRES	WG	RESEL
COST	1.			
ACRES	0.02	1.		
WG	0.29	−0.06	1.	
RESEL	−0.70	0.70	−0.24	1.

on pasture/range land. However, previous studies are not in concert with respect to the influence of land use attributes. Drake (1992), for example, reports the highest priority on forested land and the lowest on cropland.

The "warm glow" effect has a negative mean, which is evidence to more "yea-saying" than "nay-saying." The mean of the effect is very large, exceeding marginal prices of all favorable attributes, combined, i.e. respondents seem to have derived more utility from the fact of voting than the actual farmland protection they voted for. However, like all other random coefficients, it also has a large variance, i.e. there was quite diverse moral contribution to choices made by respondents.

Given the high variation of random coefficients, it is quite unlikely that a constant coefficients specification could have been preferred over RCM. Nevertheless, a likelihood ratio test was conducted, with the null hypothesis being that all random coefficients had come from degenerate distributions. Specifically, $H_0 : \mathbf{U} = 0$ was tested. The $\chi^2[10]$ distributed test statistic was 806, for which the p-value would be nearly zero, so the null was soundly rejected in favor of the RCM specification.

Out of the selection equation coefficients, the negative sign on AGE implies that older people in most cases would tend not to respond, whatever the reason might be. Similarly, men would be less likely to respond, compared to women. The influence of college education on selection was found to be positive.

The estimated correlations of random coefficients provide an additional insight. The coefficients on COST and ACRES are almost uncorrelated, which implies that there is no relationship between the marginal price of a program's size and the person's marginal utility of income. This result is a little odd, since it contradicts a common belief that people with a higher income level (people having lower marginal utility of income) are the ones concerned with an aesthetically-pleasing living environment. A mild but noticeable correlation of 0.29 between the marginal utility of income and the "warm glow" effect suggests a kind of double action of the latter. Namely, this may mean that pro-environmental voters might tend to overstate the marginal price of a program's attributes.

The primary objective of the study was to estimate Georgians' WTP for conservation easement programs. According to Equation (3.2), it is distributed as the ratio of two almost uncorrelated normal random variables. The exact distribution of a ratio of two arbitrarily correlated normals is given in Hinkley (1969), and it could be used to derive the distribution of its absolute value. However, the resultant distribution function could not be possibly solved for quantiles algebraically, so samples from the WTP distribution were Monte-Carlo simulated on a 100,000 draws from the distribution of the coefficients on ACRES and COST.

According to estimated constant coefficients, a conservation program that targets prime farmland near urban areas, which is used for growing crops and vegetables, would promise the highest WTP values. Simulating the conventional Lindahl demand for it would be impractical for both economic and statistical reasons. Figure 3.1 presents a simulated demand schedule for such a program, obtained on the basis of the median voter equilibrium concept.

The demand schedule was interpolated for program sizes from 0.1 to 2.0 million acres, given the projected 2002 Georgia resident voting population of 6,065,525 million[4], using 20-node cubic splines. As one can see, the demand function has a clear-cut hyperbolic shape. The maximum price of $920 per acre occurs at the lower limit of interpolation. At a program size

---

[4]Source: US Census Bureau. Web: http://www.census.gov.

Figure 3.1: Simulated Demand Schedule
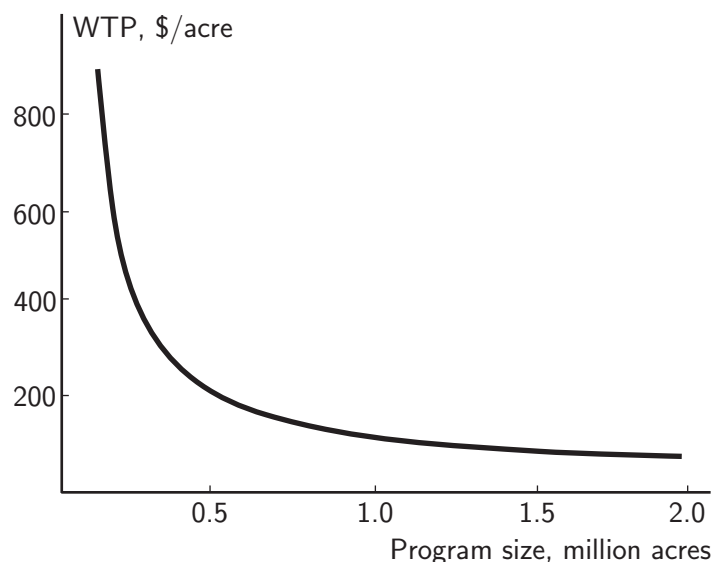
of 0.3 million acres, it drops down dramatically to \$210 per acre and then slowly decreases further, to the neighborhood of \$100 per acre.

3.8 Implications and Concluding Remarks

So far Georgia has not scored high in farmland protection through conservation easement programs. According to USDA Natural Resource Conservation Service[5], \$1,095,900 of federal funds was obligated for protecting 484 acres included in FRPP cooperative agreements in Georgia. Given the requirement of matching in-state contribution of 50% of the fair market easement value, this amounts to \$2,264 per acre to come from Georgians.

The demand schedule in Figure 3.1 makes it perfectly clear why Georgia cannot boast much progress in protecting farmland through conservation easement programs — even if one assumes a perfectly elastic supply of development rights on farmland, there can definitely be

[5]Source: FY-2002 FRPP Cooperative Agreement. Web: `http://www.nrcs.usda.gov`.

no equilibrium in the studied program size range. WTP for the smallest size of 0.1 million acres falls short of the lowest imaginable offer price by more than twice.

However, the region below 0.1 million acres is well worth further study, since the form of the demand curve promises quite high WTP values for micro-programs, sized, perhaps, to the tune of several thousand acres. This constitutes the key policy recommendation of the present study: a policy-maker concerned with welfare improvement should consider small, probably county-based conservation easement programs instead of sweeping, state-wide initiatives.

There are at least two ways to improve on the RCM model in this study. First, the present model considers marginal prices of protection quantity and an extra dollar of income to be independently and identically distributed across the population of Georgia's voters. It would be useful to investigate a possible stochastic dependence of moments of the distribution on observable socio-economic characteristics of respondents. Second, the independence of irrelevant alternatives (conditional on values of random coefficients) can be questioned. There is more and more empirical evidence showing that choices in a sequential valuation format are unlikely to be independent. Testing for dependence and, if found, accounting for it can substantively add to the reliability of value estimates.

3.9 REFERENCES

An, Y. and R. Ayala (1996, June). A mixture model of willingness to pay distributions. Duke University, Department of Economics, Working Paper 95-21. Web: `http://ideas.repec.org/p/duk/dukeec/95-21.html`.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal 100*(401), 464–477.

Bergstrom, J., B. Dillman, and J. Stoll (1985). Public environmental amenity benefits of private land: The case of prime agricultural land. *Southern Journal of Agricultural Economics 17*(1), 139–149.

Bowker, J. and D. Didychuk (1994). Estimation of the nonmarket benefits of agricultural land retention in Eastern Canada. *Agricultural and Resource Economics Review 23*(2), 218–225.

Boxall, P., W. Adamowicz, J. Swait, M. Williams, and J. Louvierre (1996). A comparison of stated preference methods for environmental valuation. *Ecological Economics 18*(3), 243–253.

Conover, W. (1999). *Practical Nonparametric Statistics* (3-rd ed.). New York, NY: John Wiley and Sons.

Drake, L. (1992). The non-market value of the Swedish agricultural landscape. *European Review of Agricultural Economics 19*, 351–364.

Eiben, A. and J. Smith (2003). *Introduction to Evolutionary Computing.* Natural Computing Series. New York, NY: Springer-Verlag.

Esseks, D. (2001). Report on 304 completed interviews in the Southeastern states. Center for Agriculture in the Environment, Public Opinion Lab, North Illinois University.

Evans, M. and T. Swartz (1995). Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. *Statistical Science 10*(3), 254–272.

Feather, P. and C. Barnard (2003). Retaining open space with purchasable development rights programs. *Review of Agricultural Economics 25*(2), 369–384.

Greene, W. (1992). A statistical model for credit scoring. New York University, Stern School of Business, Department of Economics, Working Paper 92-29. Web: `http://ideas.repec.org/p/ste/nystbu/92-29.html`.

Greene, W. (2001). Fixed and random effects in nonlinear models. New York University, Stern School of Business, Department of Economics, Working Paper 01-01. Web: `http://ideas.repec.org/p/fth/nystfi/01-01.html`.

Hanemann, W. (1984). Discrete/continuous models of consumer demand. *Econometrica 52*(3), 541–562.

Hinkley, D. (1969). On the ratio of two correlated normal variables. *Biometrika 56*(3), 635–639.

Kline, J. and D. Wichelns (1998). Measuring heterogeneous preferences for preserving farmland and open space. *Ecological Economics 26*(2), 211–224.

Lavigno, B. (2004). Is there a private market for a farmland preservation program in Georgia? Master's thesis, University of Georgia.

McFadden, D. (1994). Contingent valuation and social choice. *American Journal of Agricultural Economics 76*, 689–708.

McFadden, D. and K. Train (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics 15*(5), 447–470.

Ozdemir, S. (2003). Convergent validity of conjoint values for farmland conservation easement programs. Master's thesis, the University of Maine.

Paterson, R., K. Boyle, M. Ahearn, A. Alberini, J. Bergstrom, L. Libby, and M. Welsh (2004). Public preferences for farmland attributes in conservation easement programs. In S. Goetz, J. Shortle, and J. Bergstrom (Eds.), *Land Use Problems and Conflicts: Causes, Consequences and Solutions*. Routledge.

Rechenberg, I. (1973). *Evolutionstrategie: Optimierung Technisher Systeme nach Prinzipien des Biologischen Evolution.* Stuttgart, Germany: Fromman-Hozlboog Verlag.

Siikamaki, J. (2001). *Discrete Choice Experiments Valuing Biodiversity Convervation in Finland.* Ph. D. thesis, University of California, Davis.

Sorensen, A., R. Greene, and K. Russ (1997, March). *Farming on the Edge.* DeKalb, IL: American Farmland Trust, Center for Agriculture in the Environment, Northern Illinois University.

Train, K. (1999). Mixed logit models for recreation demand. In J. Herriges and C. Kling (Eds.), *Valuing Recreation and the Environment.* Northhampton, MA: Edward Elgar.

Train, K. (2003). *Discrete Choice Methods with Simulation.* New York, NY: Cambridge University Press.

Wynand, P. and B. van Praag (1981). The demand for deductibles in private health insurance: A probit model with sample selection. *Journal of Econometrics 17*, 229–252.

CHAPTER 4

CASE STUDY II

A PSEUDO-SEQUENTIAL CHOICE MODEL FOR VALUING MULTI-ATTRIBUTE

ENVIRONMENTAL POLICIES OR PROGRAMS IN CONTINGENT VALUATION

APPLICATIONS[1]

---

[1]Volinskiy, D., J. Bergstrom, and C. Cornwell. To be submitted to *American Journal of Agricultural Economics.*
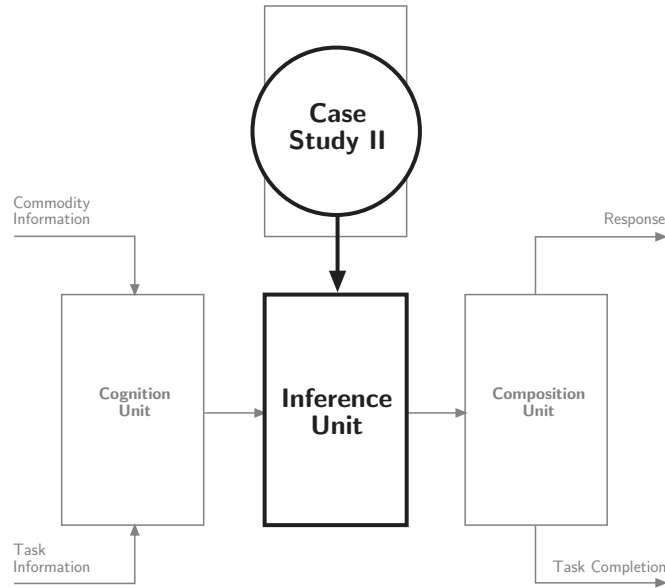
4.1 ABSTRACT

*Case Study I concluded with a recommendation to question independence of irrelevant alternatives in a sequential valuation format. Most valuation studies treat the apparent non-independence purely statistically, as a consequence of unobserved individual effects. Case Study II takes another approach. Focusing on the **Inference Unit** in the decision-making structure, Case Study II fleshes out an inferential process where any particular choice is part of a general choosing strategy of the respondent. A stochastic model is suggested, consistent with the reflexivity, transitivity and continuity axioms of utility analysis. An empirical methodology for implementing the model is suggested. An application of this model to WTP estimation for Little Tennessee River watershed ecosystem restoration is provided. Of three estimated models, the one enforcing reflexivity and transitivity and also allowing for continuity is found to have the highest in-sample predictive ability.*

4.2 Introduction and Objectives

Modern-day environmental policies or programs such as watershed ecosystem restoration are designed to improve multiple ecosystem services and consist of multiple components or parts. The valuation of such policies or programs should address the multi-dimensionality of the problem. A hybrid of the contingent valuation method (CVM) and attribute-based analysis (Holmes and Adamowicz 2003) is often used. Several related policy options are included in the survey, which are valued in a sequential manner. An example of such a sequence may be valuing a bare-bones policy first and subsequently augmenting it with more attributes or higher levels of the already included ones, building up to the most comprehensive package of management actions.

When multiple items[2] are valued using the dichotomous choice format, a binary discrete-response data set with a sequence of observations per individual is generated. The sequential nature of the choice gives rise to concerns that the probability of observing the choice outcome for a particular policy option may depend on observable or unobservable components of other choice options in the survey (Holmes and Boyle 2005). The dependence due to anchoring and framing effects, both related to the monetary bid, received some attention (Herriges and Shogren 1996; DeShazo 2002). Other causes have not been studied much. Giraud et al. (1999) provide evidence of sequencing and instrument context effect in sequential valuation. Holmes and Boyle (2005) explain the dependence by generalizing the notion of anchoring to include anchoring to the context of a specific valuation question, i.e. information in other choice sets.

In this paper, we argue that, should the econometric investigator chooses to adhere to canons of utility theory, choices in the sequential valuation format cannot be considered independent if they are conditioned on concurrent observables only. Further, a particular composition of the entire sequence entails a particular pattern of their dependence.

---

[2]Since our analysis essentially applies to valuing any composite non-market goods, the terms "commodity," "good," "policy," and "program" are used interchangeably.

The remainder of the paper is organized as follows. A general conceptual valuation model is developed in Section 4.3. On the grounds of dynamic consistency we argue that, as long as the commodity information the respondent possesses remains unchanged, the exact, albeit unobservable, utility levels attainable at all restoration programs involved should be thought of as the same throughout the valuation process. This conjecture leads to the equivalence of the sequential and simultaneous elicitation formats and makes the model consistent with the utility reflexivity and transitivity axioms.

We further build on this reasoning and posit that, for the utility continuity axiom to be maintained, the degree of dependence between the utility shocks in any pair of items should increase as the items get closer together attribute-wise. This constitutes the main research hypothesis of the study.

In Section 4.4, we introduce the specifics of survey data for the Little Tennessee River empirical application. We provide several alternative stochastic specifications for our valuation model. Model parameters estimated by maximum likelihood are presented and discussed. We discuss empirical evidence in support of the continuity hypothesis. Willingness-to-pay (WTP) values for restoration program components based on median voter equilibrium are presented and compared to the results from an earlier CVM study with the same data.

The paper concludes in Section 4.5 by discussing the ability of our model to produce economically and statistically valid welfare change estimates from data generated by the sequential, multiple valuation question per respondent format. Further enhancements are also discussed.

4.3 Conceptual Model

Consider this admittedly contrived example. The investigator is interested in comparing the agent's preferences between three hypothetical states of the world yielding deterministic utility levels v_0, v_1 and v_2. One way to elicit the preferences is to let the agent pick the preferred state from all possible pairs. An alternative is to ask the agent to rank the three

states at once. From an economic theory perspective, choosing the simultaneous format over the sequential or vice versa is immaterial as long as preferences remain unchanged. But it is not so when it comes to empirical modelling. If choices are arranged in pairs, the following random utility model (RUM) arises:

$$u_{jt} = v_j + \varepsilon_{jt} \tag{4.1}$$

$$u_{kt} = v_k + \varepsilon_{kt}$$

where (j, k) are $(1, 0), (2, 0), (2, 1)$ respectively for $t = 1, 2, 3$; v are deterministic components of the respective random utility levels and ε are utility shocks.

Some stochastic specifications may assume that utility shocks are independently and identically distributed; an error components model would account for unobserved heterogeneity of individuals. Whatever model is used, however, it will operate implicitly assuming there are six random quantities involved and there are eight possible outcomes, of which two are intransitive. Evidently, reflexivity is also violated.

Meanwhile, there are only three random quantities in the simultaneous ranking format, and there are six transitive outcomes. Thus, results from a simultaneous model will differ systematically and to an unknown extent from those coming from a sequential model even if the deterministic parts are identically specified. To make things worse, if the study aims to address a possible instrument format effect, the investigator may erroneously conclude such an effect exists while in reality it does not.

Following this line of reasoning, the investigator will probably opt for the simultaneous format even though the necessity for the respondent to consider multiple options at once may bring about accuracy concerns. Questions will remain. Is it possible to attain an equivalent representation of both formats? Should intransitivity be excluded? And, more in general, is there a way to reconcile rationality axioms that the traditional utility theory imposes on the decision-maker's preferences within the stochastic setting of a stated-preference experiment?

Let us consider a T-period sequential binary choice model. A utility maximizer i chooses, at each period t, $t = 1 \ldots T$, in a sequence between two states of the world. These states are a

period/individual-specific "alternative" (a particular environmental policy) and a no-action baseline policy, the "status quo," with the corresponding utility levels:

$$u_{it} = v_{it} + \varepsilon_{it} \tag{4.2}$$

$$\tilde{u}_{it} = \tilde{\varepsilon}_{it}$$

where $v_{it} = v(\boldsymbol{x}_{it})$ is the deterministic utility of the alternative with attributes $\boldsymbol{x}_{it}$, the deterministic utility of the status quo is zero, and $(\varepsilon_{it}, \tilde{\varepsilon}_{it})$ are the respective error terms. The model implies the following marginal probabilities of choice outcomes:

$$\Pr(u_{it} > \tilde{u}_{it}) = \Pr(v_{it} > \tilde{\varepsilon}_{it} - \varepsilon_{it}) = F_{\tilde{\varepsilon}_{it} - \varepsilon_{it}}(v_{it}) \tag{4.3}$$

where $F_{\tilde{\varepsilon}_{it} - \varepsilon_{it}}$ is the distribution function of the difference of utility shocks at time t.

The standard practice is to use $2T$ independently and identically distributed (i.i.d.) errors (Hoehn 1991). As already mentioned, this leads to the non-equivalence of elicitation formats and potential problems with transitivity. In their comprehensive review of statistical methods with CVM data, Hanemann and Kanninen (1999) consider a model where $\tilde{\varepsilon}_{it} \equiv \varepsilon_{i0}$, $\forall t$. This condition means that an unobserved utility level of the "status quo" state, u_{i0}, is the same no matter where in the survey this state is invoked. To be consistent, one can extend this logic to all states. If $\boldsymbol{x}_{is}$ and $\boldsymbol{x}_{it}$ are identical, then:

$$\Pr(u_{it} > u_{i0} | u_{is} > u_{i0}) = 1, \ s \neq t \tag{4.4}$$

It is apparent that, if all alternative states are different, this specification restricts the number of latent random quantities to $T + 1$ and statistically forces the equivalence of the sequential and parallel choice representations from our earlier example. We will term this model the *pseudo-sequential choice* to emphasize its atemporal nature. The study (ibid.) does not put forward any justification for this restriction. Indeed, what are the reasons why one should consider imposing it?

We begin with presenting our understanding of the roles played by the deterministic and stochastic components of the structural RUM in Equation (4.2). Hanemann (1984a) provides the following definition for a generic RUM:

> A random utility model arises when one assumes that, although a consumer's utility
> function is deterministic for him, it contains some components which are unobservable
> to the econometric investigator and are treated by the investigator as random variables.

An immediate implication of this definition is that there is conceptually only one source of uncertainty in the model and it is the investigator's uncertainty. The respondent consistently applies a deterministic yet unknown decision rule throughout the elicitation process. Under the general guidance of economic theory the investigator subjectively formulates a specification of the respondent's deterministic utility. There is no theory to substantiate any parametric assumption with respect to errors in the model, and both deterministic and stochastic parts need to be functionally specified before the actual survey data are incorporated into the model[3]. If so, then a distributional assumption with respect to utility shocks, whatever it turns out to be, determines the subjective rule the investigator will apply to specify the likelihood of a choice outcome for any parameter values in both parts of the model. That is, conditional on v_{it}, the investigator presents her subjective odds that $u_{it} > \tilde{u}_{it}$ by making a parametric assumption about the distributions of ε_{it} and $\tilde{\varepsilon}_{it}$.

Because the investigator has made subjective judgments with respect to both parts of the RUM, she has set the modelling rules which must be followed. This leads to determinism in her understanding of the respondent's behavior. Thus, we predicate non-volatility of both agent and investigator in their decision making. This determines dynamic consistency of the parties. Theories of rational dynamic choice generally uphold dynamic consistency. The agent should be dynamically consistent in her actions, so that, if the agent's present "self" embarks on a course of action, all later "selves" should abide by that commitment (McClennen 1990). Dynamic consistency has a timing invariance property: a sequential choice problem and

---

[3]In a number of studies, researchers attempted to obviate the issue of supplying parametric specifications for either or both systematic and random terms by using semiparametric (Klein and Spady 1993) or nonparametric approaches (Matzkin 1992; Matzkin 1993). The resultant models, however, either replace the problem of parametric specification with that of selecting a kernel and its parameters or are unsuitable to obtain a welfare change estimate.

a planned choice problem should be equivalent to the agent, given they are strategically equivalent[4].

Participants of CVM experiments are likely to have no experience with programs or policies to be valued. A number of empirical studies found that different ways of supplying commodity-related information or different amounts of information supplied led to in a significant variation in valuation results (Bergstrom et al. 1989). The time dimension and sequencing of choice sets can only be reasonably omitted in situations where information about the programs is supplied to respondents strictly prior to elicitation, and no additional information is given in between elicitation questions.

The pseudo-sequential structure of a choice model makes the latter consistent with the transitivity axiom, so that the consumer is supposed to be able to order her preferences amongst policy alternatives in a consistent manner. The essence of this model is that, if $\boldsymbol{x}_s = \boldsymbol{x}_t$[5] for Program$_s$ and Program$_t$, then $\varepsilon_s = \varepsilon_t$; that is, the respective utility shocks are perfectly dependent. But suppose there is an infinitesimally small difference between $\boldsymbol{x}_s$ and $\boldsymbol{x}_t$. Now one deals with two random quantities, ε_s and ε_t. But the utility continuity axiom, however, asserts that in this case the departure from perfect dependence should be small as well. Loosely stated, the principle of continuity postulates that any two states which are infinitely close cannot be far apart in terms of their respective utility levels. Considering environmental policies as bundles of services to the consumer, continuity allows for the possibility of substitution between policy components, which, in turn, permits comparing the relative importance of these components.

---

[4]Experimental revealed-preference studies conducted by behavioral scientists do not appear to have come up with definitive results with respect to timing invariance (Read et al. 2001). Read and Loewenstein (1995) introduced the term "diversification bias," referring to a demonstrated excess variety in items selected in the simultaneous design. Read, Loewenstein, and Rabin (1999) argue that simultaneous choice enables agents to diversify their assets to reduce the overall risk, thus giving preference to the simultaneous choice. In these and other studies, reviewed by Read et al. (2001) agents showed some psychological phenomena which admit various interpretations. The issue of an empirical validity of timing invariance largely remains a moot point.

[5]The observation subscript i will be dropped to simplify notation.

Empirical evidence seems to support this conjecture of close dependence. If a small change in the scope of a commodity or an attribute of its would lead to considerable change in utility, CVM responses would demonstrate hypersensitivity to such changes. However, more studies seem to be suffering from abnormally low sensitivity, which has lead to a broad scholarly interest to the phenomena of embedding and scope insensitivity.

Unlike real-life choices, policy alternatives in CVM studies are made distinct in a number of key attributes which are communicated to respondents. Respondents determine their preferences on the basis of what they have been told and/or shown about the choice options. If the attributes of two options are the same, there can be nothing else to distinguish them. It also follows that, provided the deterministic utility is not badly misspecified, the errors are likely to reflect some unknown joint effect of option attributes that could not be modelled within the deterministic utility specification. An example of such an effect is the agent's overall subjective perception of a policy alternative.

Assume the investigator has a measure of dissimilarity between two states, $\lambda_{st} = \lambda(\boldsymbol{x}_s, \boldsymbol{x}_t)$, increasing as dissimilarity grows. If pairs of choices can be compared on the basis of that measure, continuity would imply that unobserved utility terms for a pair of adjacent options are, on average, closer to each other compared to either out of the pair and a non-neighboring third. For example, if, for a chosen λ, $\lambda(\boldsymbol{x}_r, \boldsymbol{x}_s) \ll \lambda(\boldsymbol{x}_r, \boldsymbol{x}_t)$ and $\lambda(\boldsymbol{x}_r, \boldsymbol{x}_s) \ll \lambda(\boldsymbol{x}_s, \boldsymbol{x}_t)$, then one should expect that $|\varepsilon_r - \varepsilon_s| < |\varepsilon_r - \varepsilon_t|$ and $|\varepsilon_s - \varepsilon_t| < |\varepsilon_r - \varepsilon_t|$ to be (subjectively) more probable events than their respective complements.

We have assumed the existence of a dissimilarity measure and its relationship to the hypothesized distribution of error terms. An empirical study would certainly require a specification of this measure. Utility theory abstractly defines continuity by asserting that, for any bundles $\boldsymbol{x}$ and $\boldsymbol{y}$ in the consumption space $\{\boldsymbol{x}|\boldsymbol{x} \succeq \boldsymbol{y}\}$ and $\{\boldsymbol{x}|\boldsymbol{x} \preceq \boldsymbol{y}\}$ are closed sets. Except for the values of attributes and the utility levels, there is nothing to use to assess the dissimilarity between a pair of bundles.

It is contended that, in general, the choice of a dissimilarity function can only be implemented on subjective grounds. A usual distance function or the gravity formula can be suitable candidates. We will refrain from suggesting any particular specification. Instead, we will first focus on the effects of increasing the level of one attribute while keeping the others constant. If our conjecture is true, then statistical dependence between errors should decrease as the goods are made more different in at least one dimension.

This can be put in the form of a formal hypothesis. Let $\lambda_{st} \in [0,1]$ be a measure of statistical independence[6] between ε_s and ε_t, as above, and let $\triangle^k$ be the only different attribute k in $\boldsymbol{x}_s$ and $\boldsymbol{x}_t$, $\triangle^k = ||\boldsymbol{x}_s - \boldsymbol{x}_t||$, $\boldsymbol{x}_s^{-k} = \boldsymbol{x}_t^{-k}$. Then

Hypothesis. $\lambda_{st} \propto \triangle^k \left| \Im_s = \Im_t \right.$, where $\Im_s, \Im_t$ are the information the respondent has about the commodities at time s and t, respectively.

When $\triangle^k = 0$, the agent will assign the same utility level to all occurrences of the same hypothetical state — this is transitivity. The monetary bid is not an attribute of a bundle of commodities, so it does not influence whether commodities are closer together or farther apart. Therefore we will not be viewing the bid as part of $\boldsymbol{x}_t$. Finally, the role of constant information is to disallow the agent to reconsider expected experiences from the hypothetical commodities in light of any new information to come about in between choices.

Combining the major elements of our reasoning, we arrive at the following pseudo-sequential choice model:

$$
\begin{aligned}
u_{it} &= v_{it} + \varepsilon_{it} \\
u_{i0} &= \varepsilon_{i0}
\end{aligned}
\tag{4.5}
$$

where the distribution of $(\varepsilon_{i0}, \varepsilon_{i1}, \ldots, \varepsilon_{iT})$ in general may have $\binom{T+1}{2}$ parameters of dependence for all possible pairs of shocks. If the measure of dependence is bounded, for example, between zero and one, then normalization is required. One should select at least one pair,

---

[6]It can be, but is not limited to unity minus squared correlation.

either actual or imaginary, for which no dependence is allowed. We suggest setting dependence to zero for all $(\boldsymbol{x}_t, \boldsymbol{0})$ pairs, since the baseline option is by default most different from the rest of the policies. This results in the availability of $\binom{T}{2}$ dependence parameters.

In order to estimate the model in Equation (4.5) by maximum likelihood, one should be able to obtain probabilities of all 2^T choice outcomes. Getting the outcome probabilities for a two-period model is fairly straightforward:

$$
\begin{aligned}
\Pr(u_{i1} < u_{i0}, u_{i2} < u_{i0}) =&\Pr(u_{i0} = \max(u_{i0}, u_{i1}, u_{i2})) && (4.6)\\
\Pr(u_{i1} > u_{i0}, u_{i2} < u_{i0}) =&\Pr(u_{i2} < u_{i0}) - \Pr(u_{i0} = \max(u_{i0}, u_{i1}, u_{i2}))\\
\Pr(u_{i1} < u_{i0}, u_{i2} > u_{i0}) =&\Pr(u_{i1} < u_{i0}) - \Pr(u_{i0} = \max(u_{i0}, u_{i1}, u_{i2}))\\
\Pr(u_{i1} > u_{i0}, u_{i2} > u_{i0}) =&1 - \Pr(u_{i1} < u_{i0}, u_{i2} < u_{i0})\\
&-\Pr(u_{i1} > u_{i0}, u_{i2} < u_{i0})\\
&-\Pr(u_{i1} < u_{i0}, u_{i2} > u_{i0})
\end{aligned}
$$

For larger numbers of periods, outcome probability formulae become unwieldy, which implies having quite a complex likelihood function. While algebraic expressions for outcome probabilities grow prohibitively complex, the computation of those is easily automated, using the fact that the outcome probability for a subset of $\mathbf{Y}_i = \{y_{i1} = 1[u_{i1} > u_{i0}], y_{i2} = 1[u_{i2} > u_{i0}], \ldots, y_{i2} = 1[u_{iT} > u_{i0}]\}$ can be expressed as a sum of probabilities of the mutually exclusive joint events that constitute it. All that is needed is a facility to calculate $\Pr(u_{i0} = \max(u_{i0}, \boldsymbol{u}_{i,\mathcal{A}}))$, where $\boldsymbol{u}_{i,\mathcal{A}}$ are the utility levels of options in the subset $\mathcal{A}$ of indices $\{1, 2, \ldots, T\}$.

The objective is to solve a linear system $\mathbf{A}\mathbf{p} = \mathbf{b}$ of 2^T equations. Let $\mathbf{T}_r$ be a set of $\binom{T}{r}$ unique ordered combinations of subscripts, $\mathbf{t}_{rs}$, in $\{1, 2, \ldots, T\}$. The algorithm is formally presented in Algorithm 1.

In the next section, we provide specifications for both deterministic and stochastic parts of the model. We then fit several alternative specifications with actual survey data to assess the reasonability of our conjectures about an agent's rationality.

Data: $T,\mathbf{Y}_i,\ \mathbf{T}_r$

Result: p_j

begin

  list all possible 2^T outcomes for $\mathbf{Y}_i$;

  /*probabilities of outcomes are the unknowns in **p** */

  arrange all $\mathbf{t}_{rs}$ from all $\mathbf{T}_r,\ r = 1, 2, \ldots\ T$ in an array of sets $\mathcal{A}$;

  /*$\mathcal{A}$ will then have $\sum_{r=1}^{T} \binom{T}{r} = 2^T - 1$ elements $\mathcal{A}_j$ */

  /*when $T = 3$, $\mathcal{A} = \left\{\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\{1,2,3\}\right\}$ */

  foreach $\mathcal{A}_j$ **do**

   calculate $b_j = \Pr(\mathbf{Y}_{\mathcal{A}_j} = \mathbf{0})$;

  end

  /*For the $T = 3$ example */

  /*$b_5 = \Pr(\mathbf{Y}_{\mathcal{A}_5} = \mathbf{0}) = \Pr(y_{i1} = 0, y_{i3} = 0)$ */

  for $j = 1$ ***to*** $2^T - 1$ **do**

   for $k = 1$ ***to*** $2^T - 1$ **do**

    if $(\mathbf{Y}_{\mathcal{A}_j} = \mathbf{0})$ *event contains k-th outcome* **then**

     $a_{jk} = 1$;

    else

     $a_{jk} = 0$;

    end

    /*calculate all a_{jk} elements of **A**, except for the last row */

   end

  end

  $\mathbf{A}_{2^T,.} = \mathbf{0}$;

  $\mathbf{b}_{2^T} = 0$;

  /*put 1 in all cells of the last row of **A** */

  /*and last cell of **b** --- the sum of outcome probabilities must be one */

  solve $\mathbf{Ap} = \mathbf{b}$ for relevant outcome probability p_j with Cramer's rule;

  /*the determinant of **A** will be either 1 or -1, */

  /*which further simplifies calculations */

end

<p align="center">Algorithm 1: Obtaining Outcome Probabilities</p>

4.4 EMPIRICAL ILLUSTRATION:

ESTIMATING WTP FOR LITTLE TENNESSEE RIVER MANAGEMENT ALTERNATIVES

The Little Tennessee River watershed is located in Georgia, North Carolina, and Tennessee. The watershed encompasses 10,783 acres, including 18 rivers and streams and 26 lakes. The LTR watershed is used by logging, agriculture and mining industries; however, the aesthetically pleasing environment in the basin has brought about a tremendous increase in the population of people who visit or live within the watershed. In the last twenty years the population has doubled, leading to concerns about the future health of the watershed and the ecosystem services the watershed provides. The majority of land within the watershed is privately owned and private land use decisions have a major impact on ecosystem structure and function. For example, agricultural activities, such as watering cattle in streams, as well as housing and commercial developments along the streams and creeks influence water quality, a key parameter of ecosystem health.

The objectives of a recent CVM study by Holmes et al. (2004) were to develop and test a general methodology for valuing ecosystem services and to identify and value particular ecosystem services present in the Little Tennessee River watershed. To place a value on ecosystem services, a CVM survey instrument was designed and implemented. The present study uses the data set obtained through the above survey.

The survey followed a close-ended, single-bounded format. Valuation questions were posed in the "take it or leave it" way: "If a local county sales tax were to reduce your annual household income by $$BID$ each year for the next 10 years to support Program t, would you vote in favor of it?" 58 respondents ($N = 58$) provided complete sequences of votes in the survey.

The survey included 4 different programs ($T = 4$). Program 1 offered an overall watershed protection plan, whereby buffer strips along all small streams and creeks running into the LTR would be created. Programs 2–4 included partial restoration of the stream bank along a 20 mile stretch of the LTR, in addition to the omnipresent watershed protection plan. The

suggested scope of the restoration was 2 miles in Program 2, 4 miles in Program 3, and 6 miles for Program 4.

The computer-assisted bidding followed a simple adaptation structure. If the respondent had voted in favor of Program $\tau = 2, 3$, then the bid for Program $\tau + 1$ would have increased, otherwise $\tau + 1$ would have been offered at the same bid amount as τ.

The conditional indirect utility function we use for this study is a linear combination of weighted policy attributes and the bid:

$$v_{it} = -\gamma_i BID_{it} + \beta_{wp} wp_t + \beta_{2m} 2m_t + \beta_{4m} 4m_t + \beta_{6m} 6m_t \qquad (4.7)$$

where BID_{it} is an amount in \$100, asked from respondent i for Program$_t$, and $(wp_t, 2m_t, 4m_t, 6m_t)$ are indicators for attributes of the program. $wp_t = 1$ indicates the presence of the watershed protection plan, and $qm_t = 1$ indicates that the program provides for the restoration of a q-long stretch of the river, $q = 2, 4$, and 6 miles.

This specification admits an arbitrary dependence of utility on miles restored. To account for heterogeneity amongst respondents, we allow the coefficient on bid, γ_i, to be varying across the panel. It is assumed to follow log-Normal distribution with parameters μ and σ^2 to be estimated. Solving $v_t = 0$ for the bid value yields the compensating surplus welfare change measure for Program$_t$ (Hanemann 1984) as the ratio of the implicit price of its attributes to that of \$100 of extra income:

$$WTP_t = \frac{\beta_{wp} wp_t + \beta_{2m} 2m_t + \beta_{4m} 4m_t + \beta_{6m} 6m_t}{\gamma_i} \qquad (4.8)$$

One option for estimating the stochastic version of Equation (4.8) is a multivariate Normal distribution of utility shocks. It offers a general covariance structure and, accordingly, a full range of values for the dependence parameters, from independence to the perfect positive/negative correlation. Unfortunately, choice probabilities from a probit-type model are not closed-form expressions and must be simulated. Simulation is very computationally expensive and may result in a large variation of likelihood values, when the sample size is small.

Another candidate for estimating Equation (4.8) is a generalized extreme value (GEV) distribution. Despite being more restrictive in comparison with multivariate probit models, GEV models still allow sufficient flexibility. More importantly, GEV choice probabilities are directly computable, which substantially reduces the computational load and saves one from other problems related to simulation-assisted estimation.

Consider a GEV distribution that underlies the paired combinatorial logit (PCL) (Chu 1989):

$$F(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_J) = \exp[-G(e^{-\varepsilon_1}, e^{-\varepsilon_2}, \ldots, e^{-\varepsilon_J})] =$$

$$\exp[-G(a_1, a_2, \ldots, a_J)]$$

$$G = \sum_{k=1}^{J-1} \sum_{l=k+1}^{J} \left(a_k^{1/\lambda_{kl}} + a_l^{1/\lambda_{kl}} \right)^{\lambda_{kl}} \tag{4.9}$$

where J is the total number of options. Each $(k, l : k \neq l)$ pair of error terms in this distribution forms a nest, with the total number of nests equal to $\binom{J}{2}$ and λ_{kl} being a measure of independence for the members to the respective nest. When $\lambda_{kl} = 1$, members of the nest do not exhibit any significant covariation; when $\lambda_{kl} \to 0$, the dependence becomes perfect. The distribution provides the dependence parameters that meet our needs. Besides, if one sets $\lambda_{kl} \equiv 1, \forall k, l$, this GEV model reduces to multinomial logit (MNL).

In our case $J = T + 1$. Since the status quo option is assumed to be different from the others to the utmost extent, we restrict $\lambda_{0t} \equiv 1, \ t = 1 \ldots T$; that is, we will not allow any covariation between the error term of the status quo and those of the alternative options. This restriction conforms with a PCL identification requirement to have at least one λ set to unity. It also has a useful consequence: the model becomes the standard binary logit for any cross-section.

Using the PCL choice probability formula,

$$\Pr(u_{it} = \max(\mathbf{u}_i)) = \frac{\sum_{j \neq t} e^{v_{it}/\lambda_{tj}} \left(e^{v_{it}/\lambda_{tj}} + e^{v_{ij}/\lambda_{tj}} \right)^{\lambda_{tj}-1}}{\sum_{k=0}^{T-1} \sum_{l=k+1}^{T} \left(e^{v_{ik}/\lambda_{kl}} + e^{v_{il}/\lambda_{kl}} \right)^{\lambda_{kl}}} \tag{4.10}$$

and Algorithm 1, one can apply the regular maximum likelihood to estimate parameters in v_{it} and all λ.

Table 4.1: Three Model Specifications

Mnemonic	Description
PCL	PCL choice probabilities; λ_{st} unrestricted
MNL	PCL/MNL choice probabilities; $\lambda_{st} = 1, \ \forall s, t$
Logit	All errors are i.i.d. Extreme Value Type I

The adaptive nature of the bid generation leads to the endogeneity of BID for Programs 2 and 3. It is important to emphasize, however, that since outcome probabilities are obtained in the simultaneous choice framework, it is equivalent to conditioning the probabilities on all values of BID for a given individual, which makes the endogeneity of BID immaterial.

Table 4.1 summarizes the three versions of the model that we estimated with the specification of v given by Equation (4.7). The PCL specification applies no restrictions to the model in Equations (4.5) and (4.10); that is, the pseudo-sequential choice framework is used to ensure transitivity and 6 dependence parameters are estimated to see whether they are related to changes in the mileage of riverbank restoration in the manner hypothesized in the previous section. The MNL specification is also built on the pseudo-sequential choice framework but it excludes any dependence amongst utility shocks, so that choice probabilities are obtained from MNL. Finally, the Logit specification is simply a mixed logit, which addresses neither transitivity nor continuity. There are 8 i.i.d. Extreme Value Type I errors in this specification, 2 for each of 4 pairs of choices. Logit was chosen as a mainstream discrete-choice model. All three models reduce to binary logit for any cross-section.

Table 4.2 summarizes model parameter estimates for all specifications. Comparing the estimates, one can notice that respondents did not quite distinguish between Programs 1–3. Estimated coefficients on β_{2m} and β_{4m} are not significantly different from zero in all

Table 4.2: Estimated Coefficients

	Estimate (estimated standard deviation)		
	PCL	MNL	Logit
Watershed protection, β_{wp}	0.8973 (0.4368)**	1.0687 (0.4206)**	1.6371 (0.1537)***
2 miles of restoration, β_{2m}	−0.1078 (0.4650)	−0.1390 (0.5005)	−0.2602 (0.4743)
4 miles of restoration, β_{4m}	0.0934 (0.4876)	0.0959 (0.4434)	0.1196 (0.3755)
6 miles of restoration, β_{6m}	1.8613 (0.7908)**	1.9798 (0.5567)***	2.9450 (0.2538)***
Distribution of $\ln(\gamma)$:			
μ	1.1751 (0.5462)**	1.3838 (0.4101)***	1.3134 (0.3488)***
σ^2	4.4105 (0.5682)***	4.0044 (0.3877)***	0.7121 (0.1963)***
Predictive ability, %	20	18	16

Significance level: ***— 99%, **— 95%, *— 90%.

Table 4.3: Estimated PCL Dependence Parameters

Program	2	3	4
1	0.01	0.78	0.99
2		0.05	0.98
3			0.01

Table 4.4: Estimated WTP

Specification	WTP Quantile, $		
	25%	50%	75%
PCL:			
Programs 1–3	7	31	128
Program 4	20	86	361
MNL:			
Programs 1–3	9	25	74
Program 4	30	80	229
Logit:			
Programs 1–3	10	41	147
Program 4	34	121	452

specifications. The restoration of 6 miles of the river produces a spectacular effect. A possible cause of such an dramatic increase may be the "bet big, win big" maxim. In each management program, the survey identified category values for a set of ecosystem services, such as habitat for fish, wildlife, water purity, etc. Levels of those services were defined as "low," "moderate," or "high." While other programs featured differing service levels, Program 4 has all levels at "high." It seems to be likely that the maximum improvement was the threshold to trigger both attention and considerable spending.

In order to assess the goodness-of-fit (in-sample predictive ability) of the three models, we used the percentage of correctly predicted sequences on a 1000 draws with replacement from the sample. The results were compared to a benchmark success rate of 17%[7] attainable by indiscriminately guessing the outcome on each trial, solely based on the proportions of outcomes in the sample.

The PCL specification performed the best, marginally improving on the MNL results. At the same time, the Logit specification proved to be useless in decision-making, even falling short of the simple guessing.

Estimated λ in Table 4.3 are, at a glance, consistent with our hypothesis that the dependence between unobserved utility levels decreases as the items grow farther apart attribute-wise. $\hat{\lambda}$ for neighboring Programs are very close to zero, in other words, the respective utility shocks are highly dependent. The degree of dependence plummets to almost nothing for non-adjacent options. The dependence between the difference in miles restored and estimated λ was tested with Kendall's τ nonparametric test. The value of the statistic was 0.77, which has the p-value less than 0.01. This provides a statistical confirmation to the observed pattern.

Since the coefficient on bid is assumed to be following a log-Nornal distribution, WTP calculated according to Equation (4.8) is distributed as a weighted reciprocal of this log-Normal variate. Table 4.4 presents selected WTP quantiles for all models. We report the

---

[7]Such high a figure arises because 3 outcomes out of the possible 16 make up 68% of the sample. These are: "yes" to all programs (29%), "no" to all Programs (22%), and "yes" only to Program 4 (17%).

single WTP value for Programs 1–3, assuming the insignificant estimates of β_{2m} and β_{4m} to be zero and, thus, a difference in WTP for these Programs to be undetectable.

There are no large differences between the WTP values resulting from the three model specifications. Yet the differences of 20–50% of the value's magnitude are in no case trifling, either. Parenthetically, the obtained WTP estimates are several times higher than estimates arising from the random effects probit model by Holmes et al. (2004), while the conclusion with respect to the WTP overall super-additivity remains. The results do not quite satisfy the scope test (Arrow et al. 1993), since WTP values do start increasing until the program size reaches 6 miles of river restoration. However, the triggering effect of the maximal action package in Program 4 makes this result fairly logical. The individual demand therefore appears to be more of a step-function rather than a conventional downward-sloping schedule.

4.5 Discussion and Conclusions

Listed below are three net results from the empirical part of the study:

(a) The PCL specification that enforces transitivity and allows for continuity provides a moderately better fit, *ceteris paribus*, than others that exclude either or both continuity and transitivity.

(b) Whatever specification is used, Program 4, being the most extensive management package, has a super-additive effect on WTP.

(c) Estimated dependence parameters from the PCL specification appear to exhibit the pattern the continuity hypothesis suggests: when the difference between values of an attribute increases for two policy options, the dependence between the respective utility levels diminishes.

The immediate implication of the results (a) through (c) for WTP estimates is that a model that adheres to the principles of utility analysis is capable of providing more reliable WTP estimates both economically and statistically. Even though no dramatic differences

between estimates from different models have been found, these differences are still non-negligible and may be important for policy decisions.

Why do PCL and MNL specifications do a better job predicting choice outcomes for the Little Tennessee River data set than mixed logit? As earlier mentioned, more than half of all observations in the data set are invariable sequences of "yes" and "no" votes given for all the alternatives. Roughly 50% of respondents had not changed their mind with respect to whether or not they would want any restoration of Little Tennessee River watershed. The utility from the status quo level of the river's protection had a great deal of influence on people's choices. Knowing the respondent's vote for any given program, one could flip a coin to predict the voting outcome for another program without any valuation model. The mixed logit model totally disregards this fact. It allows the utility of the baseline state to change so that, after conditioning on the person-specific marginal utility of income, any previous or subsequent choices bear no additional information. Meanwhile, PCL and MNL specifications anchor the utility from the alternatives to the unvarying individual point of reference and thus make use of this information. These specifications provide a better control for individual heterogeneity rather then impute the series persistence to the "warm glow" or protest voting phenomena. The PCL specification goes further and reaps a reward. Based on utility continuity, it allows the utility levels from similar states to be also similar. This lets the model extract even more information from the unobserved utility components, while doing so in a manner consist with utility theory.

The authors do not intend to promote the use of paired combinatorial logit or any other particular distribution of utility shocks. Choosing such a distribution the investigator chooses her risk management technique for one thing, and a computational device, for the other. The message of our entire exercise is more general. It serves to stress the importance of specifying a stochastic CVM model in such a way that the investigator can attach theoretically-found meaning to all parameters in the empirical model. Modern statistical software allows fitting a variety of flexible probabilistic choice models. But if a particular chosen model accounts

for unobserved phenomena only mechanistically, then the researcher is left with the need for *ex post facto* interpretation of estimates. This limits the extent of quality control, since one would never know whether the observed pattern in estimates is what one should reasonably expect or it is a mere sporadic occasion. The mechanism of a dose-response statistical model ultimately reflects on the welfare estimates. If utility shocks are allowed to follow whatever process, then welfare change estimates have whatever meaning. This is definitely not what a CVM investigator would intend to produce.

Much further research needs to be done in the valuation of multiple environmental policies. A rigorous testing of the timing invariance property in CVM applications is particularly desirable. An in-depth inquiry on specifications for the stochastic interaction of utility would be instrumental for the practitioner's needs. Considering similarities between the utility space and a geographical one, a direction for search can be spatial statistical models (Besag 1975). Those models adopt a conditional probability approach, the spatial Markovity, in formulating entire spatial systems and provide holistic schemes where deterministic and stochastic components are inherently interrelated.

4.6 REFERENCES

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman (1993, January). Report of the NOAA panel on contingent valuation. *Federal Register 58*(10), 4601–4614.

Bergstrom, J., J. Stoll, and A. Randall (1989). Information effects in contingent markets. *American Journal of Agricultural Economics 71*(3), 685–691.

Besag, J. (1975). Statistical analysis of non-lattice data. *The Statistician 24*(3), 179–195.

Chu, C. (1989). A paired combinatorial logit model for travel demand analysis. In *Proceedings of Fifth World Conference on Transportation Research*, Volume 4, pp. 295–309.

DeShazo, J. (2002). Designing transactions without framing effects in iterative question formats. *Journal of Environmental Economics and Management 43*, 360–385.

Giraud, K., J. Loomis, and R. Johnson (1999). Two valuation questions in one survey: is it a recipe for sequencing and instrument context effects. *Applied Economics 31*, 957–964.

Hanemann, W. (1984a). Discrete/continuous models of consumer demand. *Econometrica 52*(3), 541–562.

Hanemann, W. (1984b). Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics 66*(3), 332–341.

Hanemann, W. and B. Kanninen (1999). The statistical analysis of discrete-response CV data. In I. Bateman and K.G. Willis (Eds.), *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*. New York, NY: Oxford University Press.

Herriges, J. and J. Shogren (1996). Starting point bias in dichotomous choice valuation with follow-up questioning. *Journal of Environmental Economics and Management 30*, 112–131.

Hoehn, J. (1991). Valuing the multidimensional impacts of environmental policy: Theory and methods. *American Journal of Agricultural Economics 73*(2), 289–299.

Holmes, T. and W. Adamowicz (2003). Attributed-based methods. In P. Champ, K. Boyle, and T. Brown (Eds.), *A Primer on Non-market Valuation*. Amsterdam, the Netherlands: Kluwer Academic Publishers.

Holmes, T., J. Bergstrom, E. Huszar, S. Kask, and F. Orr (2004). Contingent valuation, net marginal benefits, and the scale of riparian ecosystem restoration. *Ecological Economics 49*, 19–30.

Holmes, T. and K. Boyle (2005). Dinamic learning and context-dependence in sequential, attribute-based, stated-preference valuation questions. *Land Economics 81*(1), 114–126.

Klein, R. and R. Spady (1993). An efficient semiparametric estimator for discrete choice models. *Econometrica 61*(2), 387–422.

Matzkin, R. (1992). Nonparametric and distribution-free estimation of the binary choice and the threshold-crossing models. *Econometrica 60*(2), 239–270.

Matzkin, R. (1993). Fully nonparametric estimation of polychotomous choice models. *Journal of Econometrics 58*, 137–168.

McClennen, E. (1990). *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge; New York, NY: Cambridge University Press.

Read, D., G. Antonides, L. van den Ouden, and H. Trienekens (2001). What is better: simultaneous or sequential choice? *Organizational Behavior and Human Decision Processes 84*(1), 54–70.

Read, D. and G. Loewenstein (1995). Diversification bias: explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology: Applied 1*, 34–49.

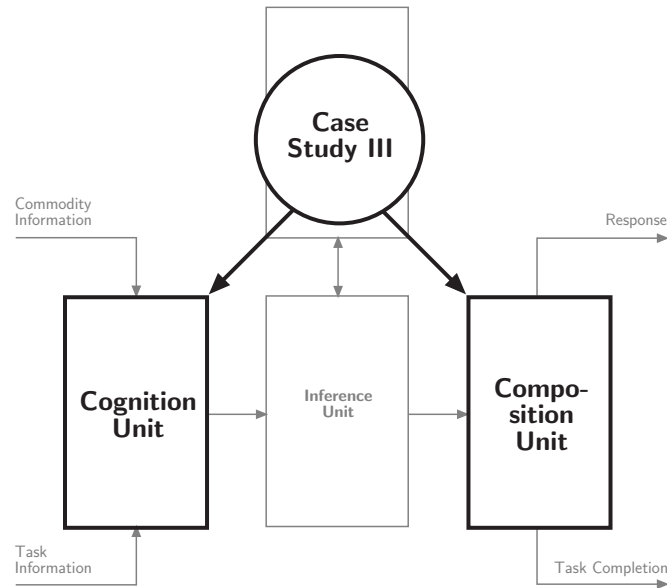Read, D., G. Loewenstein, and M. Rabin (1999). Choice bracketing. *Journal of Risk and Uncertainty 19*, 171–197.

CHAPTER 5

CASE STUDY III

VALUING MEANINGS: VALUATION OF ECOSYSTEM DAMAGE RESTORATION WITH A

FUZZY LOGIC ADVISOR[1,2]

5.1 ABSTRACT

Case Studies I and II largely ignore the existence of the **Cognition and Composition Units**. *Economic theory has no place for perception and interpretation. It does not tell how attributes, quantity, and price of a commodity are transformed into notions. Nor does it tell how vague human judgment is put into the requested elicitation format. Economic models operate with numerical and logical entities. Survey respondents judge using blurry notions which are not quantities. Case Study III looks at contingent valuation of an environmental commodity as a problem of mapping judgments which are not exactly quantifiable. The use of an expert system, a fuzzy logic advisor (FLA), to analyze people's judgements about characteristics of the commodity is discussed. A methodology of FLA-based WTP assessment and attribute analysis is provided. An application of the FLA is illustrated with valuing restoration of damage caused by the Southern Pine Beetle to public forests in the Southern Appalachians.*

5.2 Introduction and Objectives

The contingent valuation method (CVM) has both greatly benefited from and received criticism for being an interdisciplinary effort. Primarily an exercise in measuring economic concepts, a CVM experiment has to be, to no small extent, a study in cognitive psychology and semantics. McFadden (1997) emphasized that an empirical study of economic behavior could be enhanced by paying closer attention to how perceptions are formed and how they influence decision-making. The "blue ribbon" report on CVM for the National Oceanic and Atmospheric Administration (NOAA) (Arrow et al. 1993) noted that a considerable change in value estimates may come from a minute difference in a question wording as well as the number and meanings of response options.

A thorough understanding of hypothetical choice implies the investigator's understanding of what the choice options meant to the agent and what her preference actually was before being tailored to the rigid options of elicitation format. Consider a little illustration. Speaking of a forest restoration project, respondent Jane states, "I am willing to pay $10 to have 5,000 acres of forest restored." Respondent Mary states that she is not willing to pay the same amount for the same restoration project. Imagine now that we are given an ability to read their minds. We come to know that 5,000 acres vaguely meant a big forest to Jane, who happens to own a couple of acres of land. While Mary pictured the 5,000 acres as a tiny spot, thinking of vast expanses of woods she saw in Canada. We also learn that Mary is a wary spender and she was talking about a rather high amount to pay for the unusual service. What Jane in fact meant was that she would be willing to pay a "small, pocket amount of money, say $10." Finally, "I am willing to" stands for more of a "maybe yes" in Jane's vocabulary, but "I commit myself to" to Mary.

Ignoring meanings in this case equals ignoring substantial factors that caused the two respondents to make different choices. Should the CVM investigator have used verbal, qualitative categories, such as "big forest," "small amount of money," "maybe yes," the experiment would have been much more informative. However, mathematical representation of categories

of meaning is difficult. First, verbal labels have no clear boundaries. An assumption that, for example, the label "small amount of money" covers any amount in the range from \$1 to \$25 would be as arbitrary as assuming the range from \$5 to \$50.[3] Second, categories are subject to personal interpretation. As Mendel (2001) reiterates, words mean different things to different people. As a result, a particular feeling or impression cannot be put in a single category unambiguously and with certainty.

A considerable amount of research has been devoted to modelling of ambiguity of preference statements. Several authors (Ready, Navrud, and Dubourg 2001; Wang 1997; Li and Mattsson 1995; Ready, Whitehead, and Blomquist 1995) suggested models that partition the dichotomous choice format (DC) to make it more continuous and thus to accommodate preference grades. However, their empirical models still operate with clearly defined and mutually exclusive preference grades. Categories in human reasoning are neither. Mental images and emotions have a virtually continuous gradation but the survey cannot offer its respondents an infinite choice of verbal descriptions to use. Once the choice is limited to a few, people will inadvertently differ in their opinions as to what should be put in what category.

The present study considers utilizing multi-valued continuous logic, more commonly known as fuzzy logic, to obtain WTP information in a practical CVM application. The key advantage is that fuzzy set theory allows partial membership of a point from a continuous set in many discrete categories. For example, it would make it possible to classify the area of 10,000 acres as both a "small" and "large" forest to a certain degree, depending on

---

[3]Decision theories may handle imprecision in the perception of numerical stimuli and the construction of numerical responses differently from the econometric measurement error approach. Albers (2001), for example, proposes a prominence theory. The theory is based on the prominent or full-step numbers, which are integer powers of ten, their halves, and their doubles. In her discriminatory process, the decision-maker runs through the full-step numbers in decreasing order, deciding at each step whether to add, subtract, or not to use the current number to improve the precision of representation. The process stops when she reaches the boundary of her ability to judge. Albers (ibid.) observes that, for the evaluation of money, the finest perceived full interval is roughly 20% of the largest amount involved in the decision-making task. Not only does human interpretation of numbers lack cardinality, locally it is not even exactly ordinal.

how many respondents would call it so. Unlike previous research (Van Kooten, Krcmar, and Bulte 2001), we do not focus exclusively on fuzzy preference statements. A fuzzy logic advisor (FLA) that this study is about treats people's judgments about projects' sizes, attributes, monetary bids, as well as their preferences between different types of projects as fuzzy. We also developed a methodology that allows one to obtain a schedule of WTP dependence on the project's size as well as to study the influence of the project's attributes on preferences.

Our fuzzy logic-based approach to assess WTP is conceptually different from any conventional econometric model. First, the FLA does not model uncertainty as randomness. Uncertainty is considered to be deterministic. Second, the FLA is non-prescriptive. Its primary objective is to make the range of amounts that can qualify as WTP as narrow as possible. There is no single "true" WTP when inexactness is involved, and the FLA does not advocate any exact WTP values to the policy-maker. Finally, the emphasis is on a collective consensus opinion instead of obtaining inference from each respondent acting in isolation.

The objectives of the present study are:

- to develop an FLA-based methodology of WTP assessment and commodity attribute analysis to provide management feedback,

- to design a survey scheme to supply data to the FLA, and

- to use the FLA and survey data for a case of valuing restoration options for damage to public forests, caused by the Southern Pine Beetle in the Southern Appalachian region.

The rest of the paper is organized as follows. Section 5.3 introduces the working of a fuzzy logic system, along with basic notions and definitions. Bearing in mind that the general economic reader is likely to be unfamiliar with the formal theory of fuzzy sets, we favor intuition over rigor in this presentation and limit formal mathematics and terminology to a minimum. Section 5.4 illustrates the use of FLA on an example of valuing restoration options for the Southern Pine Beetle damage. Section 5.5 concludes the study with a discussion of findings and recommendations for future research.

5.3   STRUCTURE OF A FUZZY LOGIC SYSTEM

An algebraic model approximates the relationship between quantities by means of an algebraic expression. A model-free logic system uses rules instead. For the purpose of this study, a logic system is defined to be an operator which maps an input x into an output y by mapping from a collection of input sets (categories) $\mathcal{A} = \{A_{i=1}^M\}$ to that of output sets $\mathcal{C} = \{C_{j=1}^N\}$. Any rule is a conditional statement:

if x is $\boldsymbol{\mu}_{\mathcal{A}}(x)$, **then** y is $\boldsymbol{\mu}_{\mathcal{C}}(y)$

$\boldsymbol{\mu}_{\mathcal{A}}(x)$ is a vector-valued "membership" function that establishes a relationship between x and $\mathcal{A}$; $\boldsymbol{\mu}_{\mathcal{C}}(y)$ establishes a relationship between y and $\mathcal{C}$. The "if" clause is called the antecedent of the rule, while the "then" clause is its consequent. Importantly, "is" does not mean "is equal to," but it rather means "is characterized by."

Let us look first at how this kind of mapping works with conventionally defined categories, since the FLA functions in a similar manner. Term the conventional mapping device the "crisp advisor" (CA). The CA maps from a disjoint and mutually exclusive collection of sets $\mathcal{A}$ to a disjoint and mutually exclusive $\mathcal{C}$. A conventional, "crisp" set $A \subset \mathcal{A}$ on the universe of discourse Ω can be defined through an indicator membership function $\mu_A(x) = 1$ if $(x : x \in \Omega) \in A$ and $\mu_A(x) = 0$ otherwise. The same can be done with y and any of the sets in $\mathcal{C}$. Rules simplify to the following form:

if x is A_i, **then** y is C_i

Exactness means that $\boldsymbol{\mu}_{\mathcal{A}}(x)$ and $\boldsymbol{\mu}_{\mathcal{C}}(y)$ are vectors of ones and zeros for the CA. Specifically, one element in each membership vector is unity, while the rest are zeros.

To put the operation of the CA in a CVM perspective, suppose that the CVM investigator wants to know how expensive survey respondents think would be paying some amount of money for restoration of certain acreage of forest. She organizes an experiment as follows.

First, the investigator and the respondents come to an agreement with respect to two classification systems. The first would allow one to attach a verbal label, such as "large," "medium," or "small," ($\mathcal{A}$) to any restoration acreage. The second classification system would allow one to attach a verbal label, such as "large amount," "medium amount," or "small amount," ($\mathcal{C}$) to any amount of money in a certain range. Next, every survey respondent is presented with an amount of money, the same for everybody, and asked to pair each size category in $\mathcal{A}$ with a category of $\mathcal{C}$ that best describes the respondent's attitude to the money as a potential payment for the restoration. Each respondent provides as many (A_i, C_j) pairs as there are categories in $\mathcal{A}$. Votes are recorded.

The CA should perform the following operations to map from acres to dollars:

(a) Take the project size of interest x in acres and classify it as A_i according to the first classification system.

(b) Invoke (A_i, C_j) pairs that respondents provided.

(c) Determine the consensus C_j. This can be done by using the majority rule; that is, the consequent that scored most votes for the A_i is used.

(d) Finally, describe C_j as y. This can be done by averaging all money amounts in the consensus C_j, using the second classification system.

The output of this CA is an empirical step-function $y = CA(x|\text{BID})$, where BID is the bid amount used with the panel of survey respondents. Repeating the process for a progression of BID values results in a format similar to the multiple bounded discrete-choice model (Welsh and Poe 1998)[4]. This system is a policy-making device, guided by a system of *vox populi* rules. Note its model-free architecture: unlike an econometric model, this CA does not explain the nature of how x is transformed into y; it uses the results of the transformation for policy-making.

---

[4]In this multiple bounded discrete choice model, respondents are presented with an ordered sequence of dollar thresholds. For each value in the sequence, the response is elicited as one option from the list of "definitely no," "probably no," "not sure," "probably yes," and "definitely yes."

One can easily spot a fundamental problem with the CA. People have different opinions with respect to sizes of land areas and sums of money. The investigator and survey respondents would never agree in reality on any exact classification system for acreages and dollars. The cause of the problem is that the CA logic system uses the conventional dual-valued logic. Such statements as x is A_i or y is C_j must evaluate to either "true" or "false" for the classification systems to work. Fuzzy logic extends the dual-valued logic to multi-valued logic. A fuzzy statement may evaluate partly to "true" and partly to "false." While precursors to fuzzy logic have been around for a long time, Zadeh (1965) presented a formal fuzzy set theory. To date, fuzzy logic has found various applications in engineering; for example, in signal processing and automated control. Also, it can be and is used as an approximate reasoning tool in social research, especially when verbal discourse is involved (Zadeh 1975).

Fuzzy logic is built around the concept of a fuzzy set. Similarly to the crisp set, a fuzzy set A is defined through a membership function whose value can now be any real number from 0 to 1; that is, $\mu_A(x) \in [0, 1], \forall x \in \Omega$. Thus, x is allowed to have a partial membership in A, depending on its attributes.

Fuzzy sets make it possible to have statements like

x is A_1 to $\mu_{A_1}(x)$ degree, A_2 to $\mu_{A_2}(x)$ degree, ... , A_M to $\mu_{A_M}(x)$ degree

If $\sum_{i=1}^{M} \mu_{A_i}(x) = \sum_{i=1}^{M} \phi_i = 1$, then such a statement is called the "veristic" statement. The weight ϕ_i becomes a degree of truth in the statement x is A_i. The degrees of truth ϕ_i are referred to as verities.

One can "fuzzify" x (produce a fuzzy statement) for the collection $\{A_{i=1}^{M}\}$ by using verities, as done above. The reverse process is often needed, which is called the "defuzzification" of a fuzzy set or statement. Defuzzification can be performed by obtaining a central tendency measure for the fuzzy set. One of many methods available to defuzzify a fuzzy set C_j is to compute its centroid:

$$c_j = \left(\int_Y \mu_{C_j}(y)\mathrm{d}y \right)^{-1} \int_Y y\mu_{C_j}(y)\mathrm{d}y \tag{5.1}$$

where $\mu_{C_j}(y)$ is the membership function for y in the fuzzy set C_j, and Y is the set of all possible values of y.

In practice, Equation (5.1) is often discretized by replacing the integration with averaging. For example, the centroid of y is c_1 to φ_1 degree, c_2 to φ_2 degree, $\ldots$, c_N to φ_N degree is obtained as a weighted average of c_j:

$$\bar{c} = \sum_{j=1}^{N} c_j \varphi_j \tag{5.2}$$

The FLA employs the same rule-based mapping as the CA. But input and output numbers now have unique degrees of membership in all of the categories. The design we describe bellow classifies as a singleton, type-1 fuzzy logic system (Mendel 2001; Mendel et al. 1999). Here is its mechanics.

(a) Take numerical x and describe it through its membership in $A_1, A_2, \ldots, A_M$. This is done by forming M pairs (A_i, ϕ_i), where $\phi_i = \mu_{A_i}(x)/\sum_{i=1}^{M} \mu_{A_i}(x)$. The result is a veristic statement that

x is A_1 to ϕ_1 degree, A_2 to ϕ_2 degree, $\ldots$, A_M to ϕ_M degree

Let us summarize this statement as x is $\boldsymbol{\phi}_{\mathcal{A}}(x)$.

(b) Set up rules. This is done by first computing centroids c_j of all C_j according to Equation (5.1). This way y is described conditional on A_i, by forming N pairs (c_j, φ_{ij}), $\varphi_{ij} = w_{ij}/\sum_{j=1}^{N} w_{ij}$, where w_{ij} is the number of votes received by the pair (A_i, C_j). The obtained rule states that

if x is A_i, **then** y is c_1 to φ_{i1} degree, c_2 to φ_{i2} degree, $\ldots$, c_N to φ_{iN} degree

(c) Defuzzify the conditional fuzzy consequents by finding their centroids $\bar{c}_i = \sum_{j=1}^{N} c_j \varphi_{ij}$ as in Equation (5.2). This makes the fuzzy consequent crisp:

if x is A_i, **then** y is $\bar{c}_i$

Notably, this step is similar to setting the approximant to the function value at specific nodes, in a function approximation context.

(d) Combine rules (interpolate) by using the so-called fuzzy basis function (FBF) expansion:

if x is $\phi_{\mathcal{A}}(x)$, **then** $y = \sum_{i=1}^{M} \bar{c}_i \phi_i$

Both CA and FLA perform a one-to-one mapping from $\mathcal{A}$ to $\mathcal{C}$. The FLA, however, can interpolate. It maps a weighted average of elements in $\mathcal{A}$ to a unique weighted average of those in $\mathcal{C}$, solving the problem we posed in the beginning of the study: how to deal with collections of categories which are not mutually exclusive. The output of the FLA is an empirical function $y = FLA(x|\text{BID})$, which is no longer a step-function but a highly nonlinear mapping.

Note that the FLA needs one more piece of information to work. Like the CA, it needs the weighting votes from (A_i, C_j) pairs. But it also needs the membership functions $\boldsymbol{\mu}_{\mathcal{A}}(x)$ and $\boldsymbol{\mu}_{\mathcal{C}}(y)$. Clearly, the shape of the FLA output function will depend on the membership functions of the antecedents and consequents. More in general, many aspects of the FLA design will reflect on the obtained relationship. Even set unions and intersections are not fixed with fuzzy logic and can be chosen from a variety thereof. While some particular choices imply others, there are no rules as to which ones to use. This is the price to pay for the capabilities of a fuzzy logic system, and it should be taken into account when deciding whether to adopt a fuzzy logic approach. However, these assumptions are all about the inferential architecture; they do not place any restriction on the data-generating process.

The study by Van Kooten, Krcmar, and Bulte (2001) appears to be the only published to date detailed CVM application that involves fuzzy sets[5]. The authors devised a simple and elegant way to infer on WTP, using two fuzzy sets: a set labelled "Willingness to Pay" (WTP) and another labelled "Willingness Not to Pay" (WNTP). As the amount of the monetary bid

---

[5]Fujiwara, Zhang, and Okamura (2003) provided its extended and "fuzzier" version.

grows, its membership in WTP decays and its membership in WNTP grows. An intersection of WTP and WNTP characterizes a reasonable locus of WTP values. Assuming a parametric relationship between the bid and membership values for both sets, Van Kooten, Krcmar, and Bulte (ibid.) proceeded with non-linear regression to estimate the parameters and then solved the system for the WTP value at which the membership in WTP is equal to the membership in WNTP.

The valuation methodology in the present study takes another approach to using fuzzy logic. Its role is not to explain the process that maps inputs: commodity's attributes, bid etc., into an output — a pointer to WTP. The purpose of the FLA is to mimic approximate reasoning that connects input judgments about the commodity with output judgments. The sample of respondents is seen as a "black box" in terms of the true data-generating process. The only behavioral assumption is one reigning undoubted in economics: people must be willing to pay something for a good or service they want.

The next section presents a CVM mini-study on valuing two restoration strategies to deal with Southern Pine Beetle damage to public forests. In parallel, a methodology of using the discussed above FLA design as an analytical and valuation tool is presented.

5.4 Empirical Illustration:
Valuing Restoration Options for Forest Damage Caused by Southern Pine Beetle

Within this section, Section 5.4.1 discusses damage that eventuates from Southern Pine Beetle attacks and options in restoring it; Section 5.4.2 outlines the survey design; Section 5.4.3 presents the working of the FLA, covers WTP assessment and attribute analysis, and provides a commentary on their results.

5.4.1 Damage and Restoration Options

During the recent outbreak of 1999–2003, the Southern Pine Beetle killed over one million acres of native pines in five most affected states: Alabama, Kentucky, Tennessee, and the Carolinas (Merten and Nowak 2004). More than 50% of the damage was caused by Beetle activity in the Southern Appalachian region. Wilderness areas on public lands are especially prone to Beetle attacks due to the presence of older pines and the practice of not controlling the Beetle. Outbreaks in wilderness areas can spread to adjacent forests, both public and private.

The damage from the Southern Pine Beetle attacks can be restored. The modern practice is to restore a damaged pineyland to a mix of 80% of native Yellow pines — Longleaf, loblolly, pitch, Table Mountain — and 20% of hardwoods, mainly oaks. This would be as close as one can get in restoring Southern Appalachian forests to what they were described to be like by early naturalists. Two general restoration strategies can be identified. Under one strategy, damaged areas of the forest are restored to the dominant pine species in the region; that is, a project would aim to restore the pine component while creating favorable conditions for hardwood species. An alternative is to focus directly on restoring the hardwoods while creating favorable conditions for the conifers. We formalized these two strategies as the *Pines* option and the *Hardwoods* option.

The *Pines* option would proceed as follows.

- Commercially salvage some dead trees, if possible. Fell the remaining dead trees.

- Perform a light burning of the area to burn the logging debris and kill the bushes and shrubs that would compete with trees.

- Plant Yellow Pine trees that are typically found in the Southern Appalachians at wide spacing to reduce the risk of an additional Beetle attack and to allow hardwoods to grow. Watch the pines as they grow, thinning the stand whenever necessary.

The *Hardwoods* option would proceed as follows.

- Use a more intense fire to burn damaged areas in a manner of natural wildfires that would have been typical 300 years ago. This creates the seedbed conditions for Yellow Pine to germinate, kills bushes and shrubs that would compete with pines and would not harm hardwoods, such as oaks, that are tolerant of wildfires.

- Strategically plant a number of hardwood seedlings in and around the damaged areas to increase the share of hardwoods in the forest.

While both options can work equally well in restoring the visual integrity of a Southern Appalachian forest as well as reducing the public safety and fire hazards, they are distinct in a number of important aspects. These are: ecosystem naturalness, restoration success rate, economic effectiveness, and restoration time. The *Pines* option promises a higher restoration success rate, higher economic effectiveness, and quicker restoration time. The *Hardwoods* option takes longer to implement, does not generate any revenue from salvaged timber, and has an uncertain degree of success. However, it features higher overall naturalness because planting under the *Pines* option makes the forest look more uniform, and hauling salvaged trees out of the area can create an additional disturbance to the forest ecosystem.

The primary objective of the CVM experiment was therefore to conduct comparative valuation of the *Pines* option versus the *Hardwoods* option. The objective was achieved with an aid of the FLA described in Section 5.3.

5.4.2 SURVEY DESIGN

We used two groups of factors in our analysis: object variables and linguistic variables.

Object variables are the entities which survey respondents were to make judgments about. Object variables are the restoration OPTIONs (*Pines* and *Hardwoods*), monetary BID values between \$1 and \$500, and the restoration options ATTRIBUTEs: ecosystem naturalness, restoration success rate, economic effectiveness, and restoration time. Linguistic variables summarized in Table 5.1 are the entities that describe objects. These are: SIZE, EXPENSE, IMPORTANCE, and PREFERENCE. Every linguistic variable was given a numerical metric and

Table 5.1: Linguistic Variables

Mnemonic	Fuzzy categories	Numerical metric
SIZE	"Very large" "Large" "Rather large" "Rather small" "Small" "Very small"	Acres. Sizes used: 1,000; 2,000; 5,000; 10,000; 20,000; 50,000
EXPENSE	"Way too much for what I get" "Rather costly" "Reasonable amount" "Quite inexpensive" "Almost nothing for what I get"	Likert scale 1–10
IMPORTANCE	"Top priority matter" "Moderately important" "Rather unimportant" "Not important at all"	Likert scale 1–10
PREFERENCE	"I like *Hardwoods* much better than *Pines*" "I like *Hardwoods* a bit better than *Pines*" "Cannot say which one I like better" "I like *Pines* a bit better than *Hardwoods*" "I like *Pines* much better than *Hardwoods*"	Likert scale 1–10

an associated list of fuzzy categories. Values of SIZE are, naturally, acres. Since there is no continuous numerical measure for grades of expensiveness, importance, or preference, the Likert technique with a ten-point scale was used as a numerical metric with these object variables. Numerical values of the object BID and linguistic SIZE were chosen to be full-step numbers (integer powers of ten, their halves, and their doubles) (Albers 2001) to facilitate their perception by respondents.

The survey was conducted in pursuit of two objectives: (a) to obtain weighting votes for mapping between the fuzzy categories of the linguistic variables, and (b) to acquire membership information for those fuzzy categories. The survey included three major components: presentation, elicitation, and membership function acquisition.

The presentation of damage caused by the Southern Pine Beetle and the restoration options followed their description in Section 5.4.1. In the elicitation component, respondents were asked to indicate the most appropriate IMPORTANCE category for each ATTRIBUTE. Next, respondents were asked to indicate their preference between the *Pines* and *Hardwoods* options, using the categories of PREFERENCE.

The elicitation of WTP followed the structure of the multiple bounded discrete-choice format (Welsh and Poe 1998). An increased sales tax scenario was used as the payment vehicle. The question was posed as:

> For the restoration Option below, please imagine that the sales tax increase to pay for Pine Beetle damage restoration would decrease your annual disposable income by the stated amount each year for 10 years.
>
> - Project Type: [*Pines* or *Hardwoods*]
>
> - Restoration Area: [SIZE]
>
> - Reduction in Your Annual Income: [BID]
>
> For this project, what is your reaction to this reduction in your annual income that would occur each year for the next 10 years? [list of EXPENSE categories follows]

Each respondent was asked to evaluate all 9 BID amounts for each restoration option. SIZE values from the list in Table 5.1 were randomly assigned to respondents, and the SIZE value remained the same during the elicitation. A series of questions followed up to screen out respondents which might be opposed to the payment vehicle and, accordingly, vote against any increase in the sales tax by stating that any BID value is too expensive.

Note that the type of elicitation question asked is different from what one would expect in a CVM experiment. Instead of asking whether one would be willing to pay a specified amount, we asked respondents to state how expensive paying the BID amounts they think would be. Asking directly about one's WTP has a clear connotation of selling the commodity. Asking to evaluate the price is more of a seeking one's advice on what to put on the price tag. This fits the advisory function of our WTP FLA system, which is supposed to produce a consensus recommendation to the policy-maker.

Membership information was acquired for all linguistic variables by combining elements of methods by Mendel et al. (1999) and Turksen (1991). For EXPENSE, IMPORTANCE, and PREFERENCE, respondents were asked to imagine themselves in a situation were someone is using a wording and they need to place the intensity of its message on an artificial 1 to 10 (Likert) scale. Below is an instance of the EXPENSE acquisition question:

> Suppose you are discussing an amount of money that appears on the price tag of a product or service offered to you. Someone tells you it is *rather costly*. On the scale 1 to 10, where 1 means a totally negligible price and 10 means an outrageously high price, select the lowermost and uppermost numbers that you associate in your mind with this level of expense.

As a result, each respondent provided two integer scale numbers for each fuzzy category of EXPENSE, IMPORTANCE, and PREFERENCE. The process was reversed for SIZE. Respondents were presented with all SIZE values in acres, one at a time, and asked to indicate the most appropriate category for each given acreage. Each respondent thus provided one category for each SIZE acreage. Table 5.2 contains the exact wordings of all acquisition questions.

Table 5.2: Membership Acquisition Questions

Mnemonic	Question
SIZE	Suppose you hear it on the radio or TV that an area of forest in your State has been destroyed by a wildfire, insects, or disease. What would you call if the size of this area is [SIZE value]? Select one from [list of SIZE categories]
EXPENSE	Suppose you are discussing an amount of money that appears on the price tag of a product or service offered to you. Someone tells you it is: [EXPENSE category] On the scale 1 to 10, where 1 means a totally insignificant price and 10 means an outrageously high price, select the lowermost and uppermost numbers that you associate in your mind with this level of expense.
IMPORTANCE	Suppose you are discussing the objective of a governmental policy. Someone tells you that achieving this objective is: [IMPORTANCE category] On the scale 1 to 10, where 1 means the policy is absolutely unimportant and 10 means extreme importance, select the lowermost and uppermost numbers that you associate in your mind with this level of importance.
PREFERENCE	Suppose you are discussing two brands, A and B, of some consumer product. Someone says: [PREFERENCE category] On the scale 1 to 10, where 1 means your absolute preference of A and 10 means your absolute preference of B, select the lowermost and uppermost numbers that you associate in your mind with this this expression of preference.

Figure 5.1: Principal FLA Structure

(a) Valuation Design
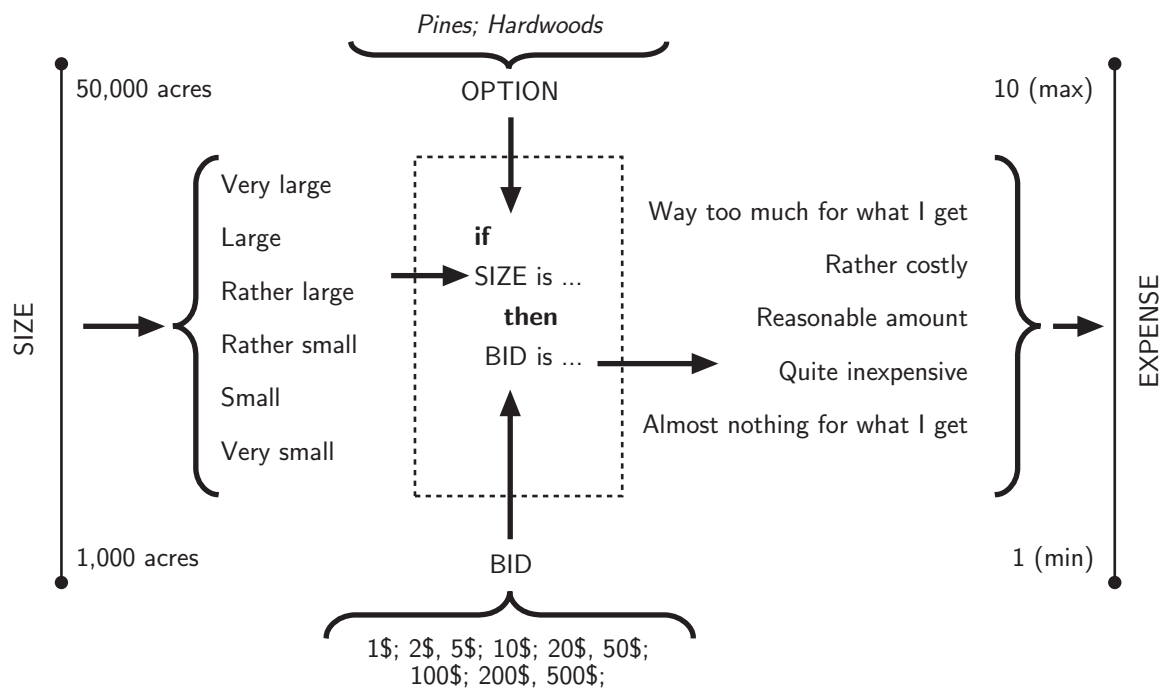
(b) Attribute Analysis Design
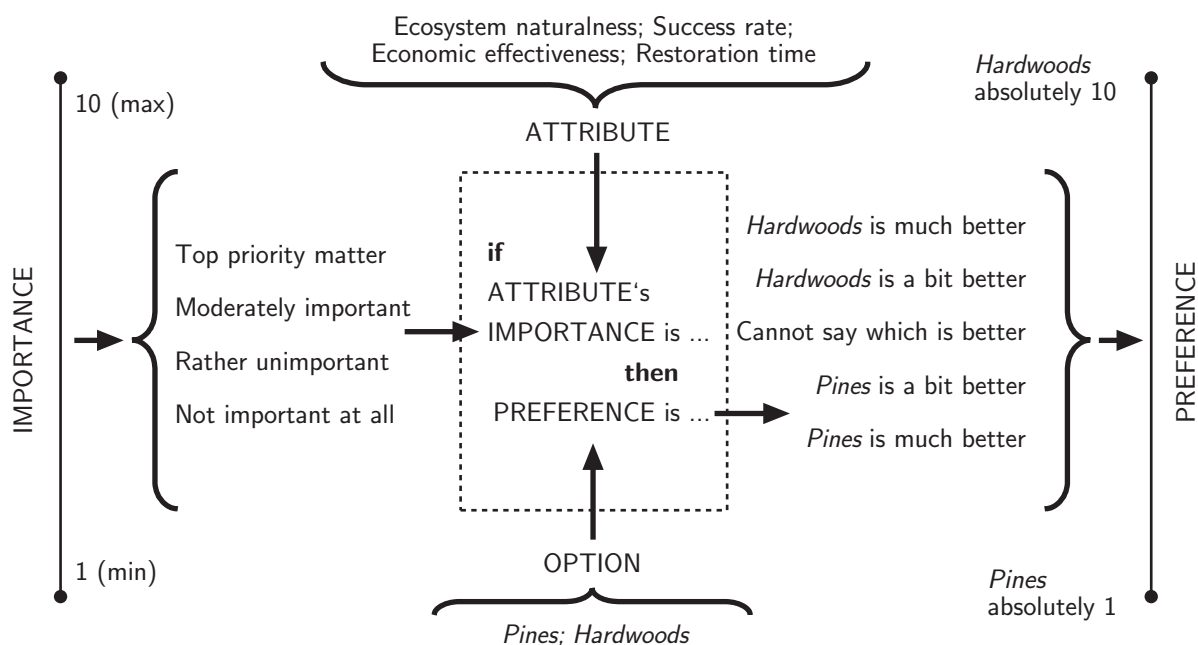
5.4.3   FLA IMPLEMENTATION AND OUTPUT

Two designs presented in Figure 5.1 carry out two main tasks in the comparative valuation of a *Pines* restoration project versus a *Hardwoods* project.

The valuation design in Figure 5.1(a) is to map from fuzzy SIZE categories to fuzzy EXPENSE categories. With reference to the general FLA design from Section 5.3, SIZE categories are the input sets $\mathcal{A}$, and EXPENSE categories are the output sets $\mathcal{C}$. The valuation FLA uses the following (A_i, C_j) pairs:

if SIZE is SIZE$_i$, **then** BID is EXPENSE$_j$

for the progression of BID values and both restoration options.

The attribute analysis design in Figure 5.1(b) maps from fuzzy IMPORTANCE categories to fuzzy PREFERENCE categories. IMPORTANCE categories are now the input sets $\mathcal{A}$ for this design, and PREFERENCE categories are the output sets $\mathcal{C}$. The attribute analysis FLA uses the following (A_i, C_j) pairs:

if ATTRIBUTE is IMPORTANCE$_i$, **then** OPTIONs are PREFERENCE$_j$

for all four ATTRIBUTE values, where the preferential relationship is that between the *Pines* option and the *Hardwoods* option.

As common in engineering applications, triangular and piece-wise linear membership functions were constructed. In Figure 5.2, the functions on the left and right are piece-wise linear (two shape parameters); the one in the middle is triangular (three shape parameters). For all non-boundary categories of SIZE, the triangular function was constructed by using the averages of project sizes in the respective categories as the apex m, the minimum size in the category as the left-end base point a, and the maximum as the left-end base b. The piece-wise components were built in the same manner, discarding the minimum or the maximum, respectively. For EXPENSE, IMPORTANCE, and PREFERENCE, a was set to the average of Likert scale points for the lower category limit minus one standard deviation of these points,

Figure 5.2: Triangular Membership Functions
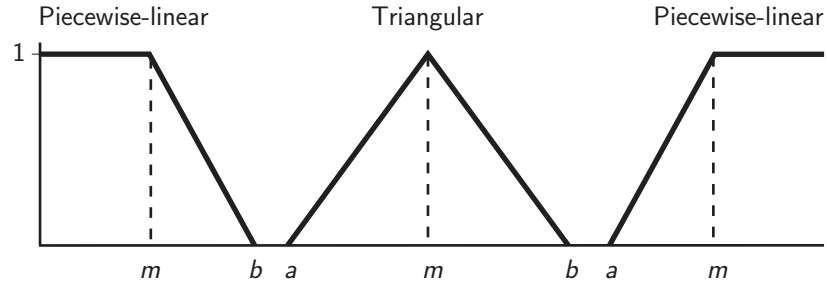
b was that for the upper category limit plus one standard deviation, and m was set equal to the overall average of points for both limits.

The survey was designed and deployed as a Web-based application by an independent market research company. Administered to members of general population in western North Carolina and eastern Tennessee, it generated 219 responses. Major descriptive statistics of the sample can be found in Table 5.3. As already mentioned, a series of questions was ask to every respondent in a effort to detect respondents who had a negative attitude towards the payment vehicle. Only votes of those respondents who explicitly approved of the vehicle were used for the valuation FLA (Figure 5.1(a)), which led to having 125 observational units available for WTP assessment.

Raw FLA output for both *Pines* option and *Hardwoods* option is shown in Figure 5.3. Two characteristics of the schedules are especially noteworthy: their virtual flatness beyond 5,000 acres and overall similarity between the two options. Bid schedules become practically flat in a region between 5 and 10 thousand acres restored; for that reason, Figure 5.3 only displays the schedules up to the size of 10 thousand acres. A normal consequence of monotonicity or non-satiation would be each bid amount becoming ever less expensive as the project size grows. For large project sizes in the obtained schedules, however, the attitude towards any bid amount remains the same regardless of restoration acreage. Three possible reasons

Table 5.3: Sample Statistics

Total	219 respondents, where:
GENDER	53% female, 47% male
AGE	average 49.7 years
	std. deviation 14.3 years
INCOME	average 59,800$
	std. deviation 39,900$
SCHOOLING	0.5% Jr. High
	18.1% High school
	47.5% College/technical
	33.9% Graduate/professional

Figure 5.3: Raw Output of Valuation FLA

(a) *Pines* Option
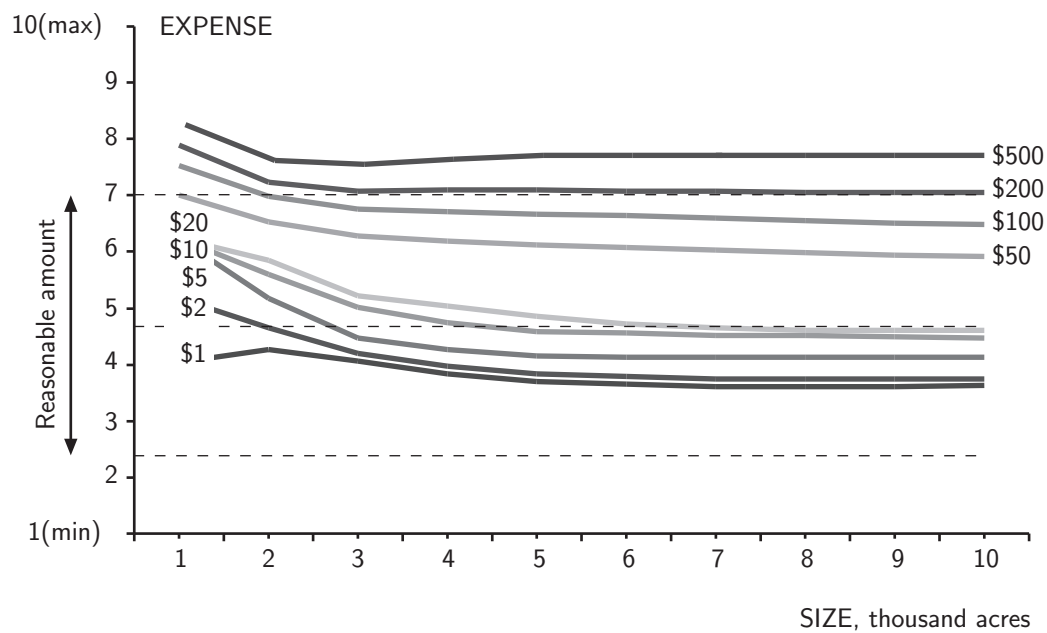
(b) *Hardwoods* Option
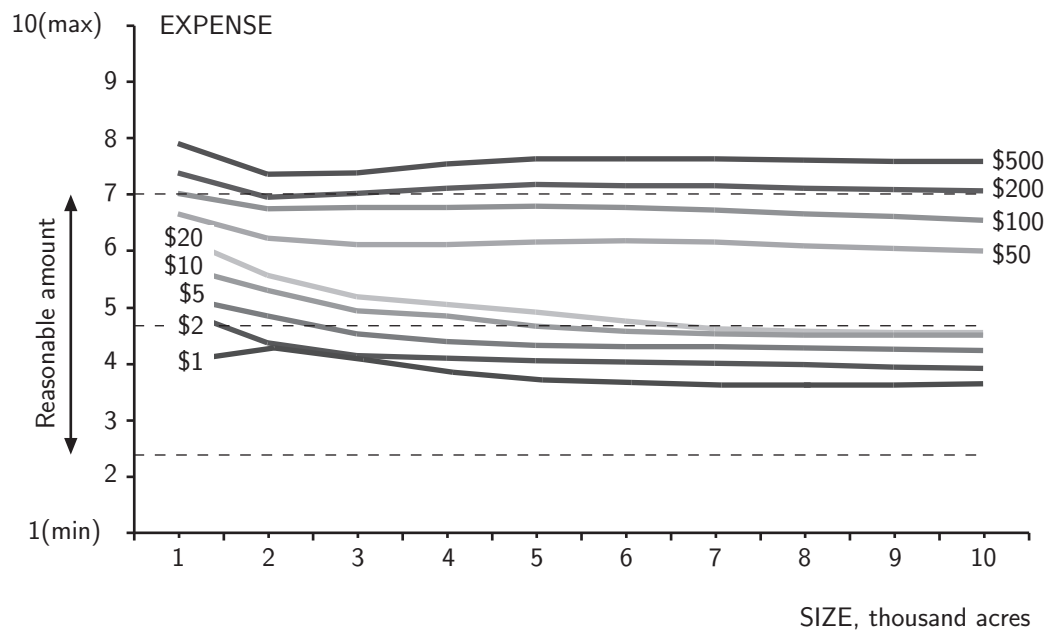
for this occurrence were hypothesized: (a) embedding effect, (b) poor perception of larger acreage sizes, and (c) a budget constraint.

Embedding generally refers to a respondent's failure to take into account the commodity size or scope when making a willingness-to-pay decision. As a result, even though a smaller commodity is embedded into more comprehensive one, value estimates appear to be the same for both. However, the use of qualitatively defined categories on the quantitative commodity size and/or attributes lowers the risk of an embedding. For instance, if, of two integers q_1 and q_2, q_1 is much greater than q_2, then the quantity q_1 of the commodity naturally includes its arbitrary q_2 units. However, a "large quantity" and "small quantity" may overlap but they are quite different categories.

Perception of numerical sizes does not seem to be a likely reason, either. Categories for larger project sizes are definitely wider[6], which reflects the logarithmic evaluation of numerical stimuli (Albers 2001). Hence their considerable overlapping. However, it cannot be concluded that respondents showed a total lack of coherence in labelling larger project sizes. The categories come in proper order, neither category is subsumed by another one, and their crests (average values) are considerably spaced. Therefore, there seems to be no dramatic problems with the perception of large numbers.

This leaves a budget constraint as a candidate explanation for the weak scope sensitivity. Becker (1965) proposed a theory of utility maximization within groups of related commodities, which includes budget pre-allocation for such groups. This might have been the case with survey respondents. Some evidence to this conjecture comes from observing that all bid amounts up to $20 appear to have been assessed as a little cheap, while bids starting at $50 seem to be slightly excessive. If respondents' budgets allocated to general expense purposes typically tend to fall between $20 and $50, this can cause the observed separation of bid amounts into two large groups: those below the budget allotted and those above it.

---

[6]Several categories with the same lower and upper bases but different tops produce the noticeable "bottleneck" shape of the schedules.

The similarity between the sets of bid schedules for the two restoration options apparently arises because the options are both state-of-the-art. Their attribute sets balance each other and, as a result, no option dominates another one. More comments on this fact are given below, in our remarks on the attribute analysis.

The information in Figure 5.3 is useful but unsuitable for decision-making purposes. The artificial EXPENSE dimension should be reduced. We remove this dimension using a maximum possibility approach. The peak of the "reasonable amount" fuzzy set (whose lower and upper bases and the peak are shown on bid schedules in Figure 5.3) indicates the level of expense which corresponds to the most possible reasonable expense. This is the closest proxy to the notion of WTP this FLA allows. Accordingly, we obtained the intersections of bid schedules with the peak of the set and re-mapped the obtained points into a project size versus WTP space.

The obtained WTP (inverse) schedules are presented in Figure 5.4. As could be expected from the similar bid schedules structure, WTP schedules for both restoration options look very similar. The WTP schedule for the *Pines* option starts at a \$1/year per thousand acres WTP for restoration of 1,950 acres and grows up to \$3/year per thousand acres for a 6,500 acres restored. The WTP schedule for the *Hardwoods* option goes from \$1 per thousand acres for restoration of 1,500 acres to \$3/year per thousand acres at a project size of 6,500 acres. Since the flatness of bid schedules precluded further interpolation, we assume that the relative WTP tops at around \$3/year per thousand acres and gradually decreases from there for larger project sizes to reach some \$0.1 per thousand acres for a restoration of 75,000 acres (Reaves, Kramer, and Holmes 1999). Overall, the obtained absolute WTP values for the Southern Pine Beetle damage restoration appear to fall into the lower part of a range of results from several forest valuation studies (Kramer, Holmes, and Haefele 2003).

Results of the attribute analysis with an FLA built according to the scheme in Figure 5.1(b) are presented in Figure 5.5. The purpose of the attribute analysis was to analyze impacts of option attributes on preference between the restoration options, in order
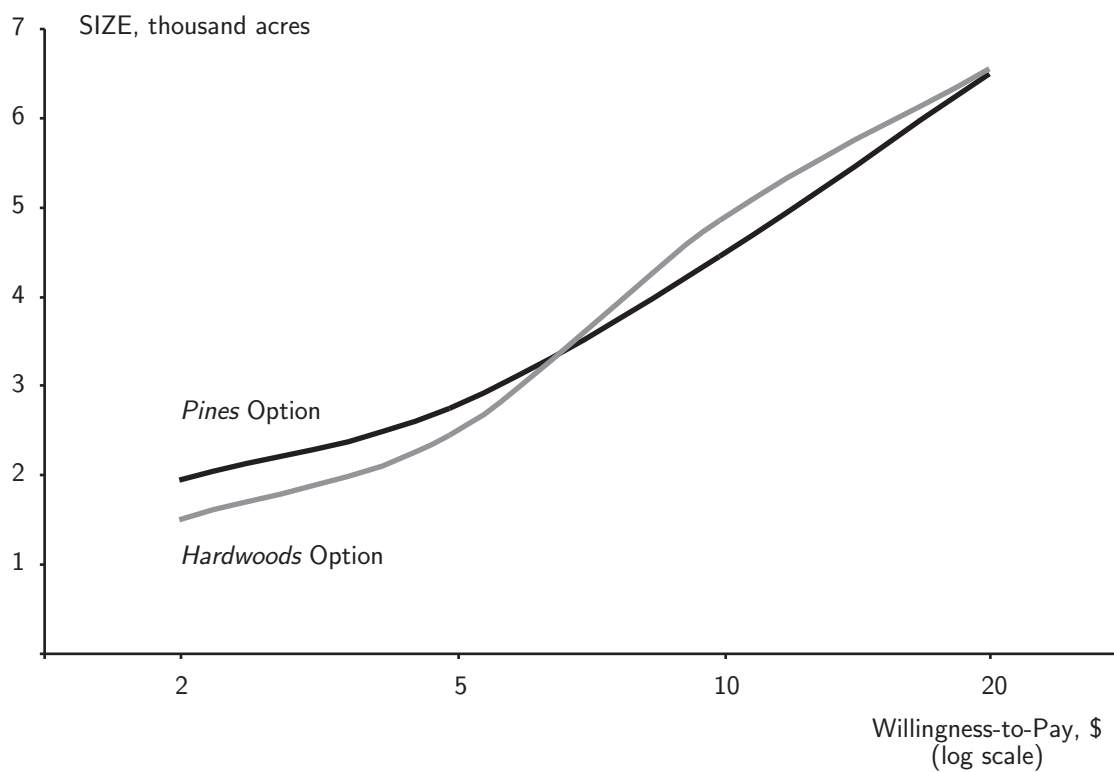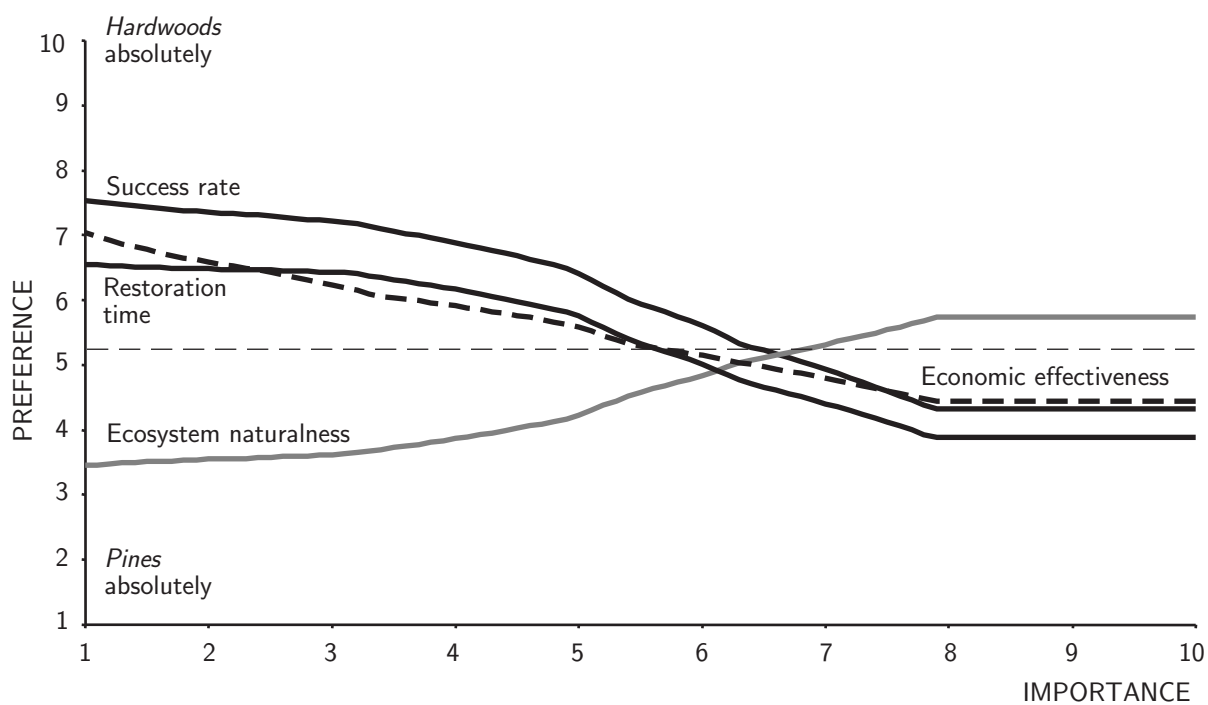
Figure 5.4: WTP Schedules

Figure 5.5: Attribute Analysis Schedules

to provide management feedback and recommend measures to increase public acceptance of a restoration project.

Survey respondents well received and meaningfully processed information on restoration options attributes, since the schedules have proper slopes and are relatively monotone. The schedules of the *Pines* option attributes: the success rate, economic effectiveness, and restoration time slope downwards. This implies that a growing importance of these attributes would make people prefer the *Pine* option over *Hardwoods* in case the importance attributed to ecosystem naturalness is not high. Conversely, a respondent would rather prefer the *Hardwoods* option if the importance of ecosystem naturalness dominates the rest of the attributes. The schedules are nearly symmetric with respect to the virtual indifference line at 5.25 values of the scale, which is the crest of the "cannot say which one I like better" fuzzy set. The implication of this symmetry is that a respondent would hardly give preference to one restoration option in case she finds all attributes to be moderately important.

Since neither restoration option turned out to provide a higher WTP, we hypothesize that ecosystem naturalness has the same weight in one's decision-making as the success rate, economic effectiveness, and restoration time, combined. Of the latter, economic effectiveness and restoration time generate a slightly higher response in terms of option preference because these two cross the indifference line a little sooner than does success rate.

We conclude this mini-study with a recommendation to consider moderately sized projects for the restoration of damage caused by the Southern Pine Beetle in the Southern Appalachian region. The optimum size for such a project would be in the range between 5,000 and 10,000 acres. People's WTP for projects in this range is estimated to be around $3/year per thousand acres. Either of the restoration options described in this study can reasonably be used. Improving the economic effectiveness and restoration time for a project adhering to active restoration of the pine component can help improve its acceptance and increase its value amongst general population of the region.

5.5 Concluding Remarks

Prominent critics of contingent valuation, Diamond and Hausman (1994) asked: "Is some number better than no number?" Their message was to avoid involving economics in determining non-use values, which should better be arrived at through a political process. Curiously perhaps, the manner in which the question was posed implies a dual-valued logic. Between some number and no number lies information. The fuzzy logic advisor discussed in this paper helps extract this information. It imitates a political process through which a board of public representatives would come to consensual WTP values. Along the way, substantial economic knowledge is obtained.

The FLA design used in this study has a significant drawback. While taking care of inexact judgments, it ignores people's uncertainty about what judgment to make. One direction to improve the FLA is an upgrade to type-2 fuzzy logic. Type-2 fuzzy logic allows handling uncertainties about the meanings of words by modelling the uncertainties. This is achieved by making the boundaries of membership functions blurry. A type-2 system could therefore directly address the issue that words mean different things to different people.

Another improvement could be adapting the FLA/valuation instrument design to analyze multiple commodity attributes simultaneously, in a manner similar to the multiple linear regression. A fuzzy logic system can accept both scalar and vector types of input. However, this causes the number of actual input categories to grow geometrically. If the attribute analysis FLA from the previous section were processing all four attributes at once, it would be dealing with a total of $5^4 = 625$ categories, which would require at least this many responses available. Finally, other ways to obtain the most possible locus of WTP values should be explored; for example, several bid values can be simultaneously communicated to the respondent.

Fuzzy logic systems are no magic answer to the problem of inexactness and perception phenomena in CVM. They are an alternative way to obtaining economic knowledge. Would some knowledge be better than none?

5.6 References

Albers, W. (2001). Prominence theory as a tool to model boundedly rational decisions. In G. Gigerenzer and R. Selten (Eds.), *Bounded Rationality: the Adaptive Toolbox*, Dahlem Workshop Reports. Cambridge, MA: MIT Press.

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman (1993, January). Report of the NOAA panel on contingent valuation. *Federal Register 58*(10), 4601–4614.

Becker, G. (1965). A theory of the allocation of time. *Economic Journal 75*(299), 493–517.

Diamond, P. and J. Hausman (1994). Contingent valuation: is some number better than no number? *Journal of Economic Perspectives 8*(4), 45–64.

Fujiwara, A., J. Zhang, and T. Okamura (2003). Evaluation of neighborhood street improvements based on fuzzy contingent valuation method. *Journal of the City Planning Institute of Japan 38*(3), 505–510.

Kramer, R., T. Holmes, and M. Haefele (2003). Using contingent valuation to estimate the value of forest ecosystem protection. In E. Sills and K. Abt (Eds.), *Forests in a Market Economy*. Boston, MA: Kluwer Academic Publishers.

Li, C.-Z. and L. Mattsson (1995). Discrete choice under preference uncertainty: an improved structural model for contingent valuation. *Journal of Environmental Economics and Management 28*(2), 256–269.

McFadden, D. (1997) Rationality for economists. Presented at the NSF Symposium on Eliciting Preferences. Berkeley, California.

Mendel, J. (2001). *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Upper Saddle River, NJ: Prentice Hall.

Mendel, J., S. Murphy, L. Miller, M. Martin, and N. Karnik (1999). The fuzzy logic advisor for social judgments. In L. Zadeh and J. Kacprzyk (Eds.), *Computing with Words in*

Information/Intelligent Systems, Volume 33 of *Studies in Fuzziness and Soft Computing.* New York, NY: Physica-Verlag.

Merten, P. and J. Nowak (2004, January). Overview of consequences of the SPB-killed pines in the Southern Appalachians and Cumberland Plateau. In *Proceedings of "After the Southern Pine Beetle" Workshop.* Southern Appalachian Man and the Bioshpere (SAMAB). Web: `http://samab.org/Focus/Ecosystem/SPB.html`.

Ready, R., S. Navrud, and W. Dubourg (2001). How do respondents with uncertain willingness to pay answer contingent valuation questions? *Land Economics 77*(3), 315–326.

Ready, R., J. Whitehead, and G. Blomquist (1995). Contingent valuation when respondents are ambivalent. *Journal of Environmental Economics and Management 29*(2), 181–196.

Reaves, D.W., R. Kramer, and T. Holmes (1999). Does question format matter? Valuing an endangered species. *Environmental and Resource Economics 14*, 365–383.

Turksen, I. (1991). Measurement of membership functions and their acquisition. *Fuzzy Sets and Systems 40*, 5–38.

Van Kooten, C., E. Krcmar, and E. Bulte (2001). Preference uncertainty in non-market valuation: a fuzzy approach. *American Journal of Agricultural Economics 83*(3), 487–500.

Wang, H. (1997). Treatment of "don't know" responses in contingent valuation surveys: a random valuation model. *Journal of Environmental Economics and Management 32*(2), 219–232.

Welsh, M. and G. Poe (1998). Elicitation effects in contingent valuation: Comparisons to a multiple bounded discrete choice approach. *Journal of Environmental Economics and Management 36*(2), 170–185.

Zadeh, L. (1965). Fuzzy sets. *Information and Control 8*, 338–353.

Zadeh, L. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences 8*, 199–249.

Chapter 6

Future Research Directions

Case Studies I and II employed static, path-independent models of consumer choice in a sequential choice framework of multiple commodity valuation. Whenever multiple items are valued, there has always been a concern that the respondent's choice can be affected by her previous choices; that is, path dependence takes place. Chapter I reviewed several framing-related effects that cause path-dependence, such as ordering effect and anchoring.

However, both explanations of what causes path-dependence and suggested models are largely *ad hoc* and mechanistic. Framing and prospect theory alone do not seem to be a universally applicable justification for path-dependent preferences. The author suggests to study learning phenomena as a direction for further research on sequential choice and path-dependence in stated-preference valuation.

Economic theory treats the decision-making mechanism of any respondent as fully formed by the time the valuation experiment starts. However, both commodities and choices involved are hypothetical; the respondent is very unlikely to have ever been to a situation like the setting of the experiment. It is therefore only natural to the respondent to learn about both the commodity and her preferences with respect to it right in the process of valuation. Learning thus becomes an integral part of economic decision-making (Rizzello 1999). Meanwhile, learning as a fundamental component of stated preferences has been largely ignored in valuation literature. Much research needs to be conducted to reveal learning mechanics and its possible representations in a valuation model.

More research is needed in the field of stochastic specification of valuation models and relations between the deterministic decision-rule component and the stochastic part, as discussed in Case Study II. Consumption bundles representing environmental amenities or policies are multidimensional and complex. Random structures that have these qualities should be closer investigated. A direction for search can be spatial statistical models and the associated random fields.

Extensive studying of perception modelling and the inclusion of psychometric factors into valuation models is required. The general undesirability of psychological effects in the economic valuation model led to perception effects having been termed biases or errors. The author believes this to be a fallacy. Commodity and task perceptions have a great deal of impact on choice outcomes and therefore need to enter the model as its full-fledged components.

Another promising research direction could be making survey instruments self-improving. Several noted scholars emphasized a need to model the entire response process in stated-preference valuation, not only the discriminal process. The model-free design of an fuzzy logic systems allows one to go even further and design intelligent surveys combining the valuation instrument and a processing system. In particular, running a fuzzy logic system in real-time should be considered. If results could be processed while the interview is still in progress, the participants could be asked to consider new, optimized bids and consensual WTP amounts could be verified.

On a more theoretical level, a richer set of welfare and value measures is needed. Stated-preference valuation is primarily an empirical exercise in measuring economic concepts. Economic theory does not tell why people like a particular commodity or how they come to liking it. When non-economic phenomena enter an economic model or no exactly defined economic model is used, the familiar willingness-to-pay and consumer demand lose their exact neoclassical interpretation. Concerns arise as to what is then being measured, and these concerns need to be addressed.

References

An, Y. and R. Ayala (1996, June). A mixture model of willingness to pay distributions. Duke University, Department of Economics, Working Paper 95-21. Web: `http://ideas.repec.org/p/duk/dukeec/95-21.html`.

Arrow, K. and A. Fisher (1974). Environmental preservation, uncertainty, and irreversibility. *Quarterly Journal of Economics 88*(1), 312–319.

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman (1993, January). Report of the NOAA panel on contingent valuation. *Federal Register 58*(10), 4601–4614.

Barron, J. and J. Greene (1996). Determinants of insensitivity to quantify in valuation of public goods: Contribution, warm glow, budget constraints, availability, and prominence. *Journal of Experimental Psychology: Applied 2*, 107–125.

Bergstrom, J. and J. Loomis (1999). Economic dimensions of ecosystem management. In H. Cordell and J. Bergstrom (Eds.), *Integrating Social Sciences with Ecosystem Management*. Champaign, IL: Sagamore Publishing.

Bergstrom, J. and J. Stoll (1990). An analysis of information overload with implications for survey design research. *Leisure Sciences 12*, 265–280.

Bergstrom, J., J. Stoll, and A. Randall (1990). The impact of information on environmental commodity valuation decisions. *American Journal of Agricultural Economics 72*(3), 614–621.

Bilmes, J. (1998, April). A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. International Computer Science Institute. Paper TR-97-021.

111

Blackwell, R., P. Miniard, and J. Engel (2001). *Consumer Behavior* (9th ed.). Ft. Worth, TX: Harcourt College Publishers.

Calia, P. and E. Strazzera (1999). A sample selection model for protest non-response votes in contingent valuation analyses. Intomics RePEc Archive. Paper 199905. Web: `http://inomics.com/cgi/repec?handle=RePEc:cns:cnscwp:199905`.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies 47*, 225–238.

Cummings, R., D. Brookshire, and W. Schulze (1986). *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*. Totowa, NJ: Rowman and Allanheld.

DeShazo, J. (2002). Designing transactions without framing effects in iterative question formats. *Journal of Environmental Economics and Management 43*, 360–385.

Desvouges, W., F. Johnson, R. Dunford, S. Hudson, K. Wilson, and K. Boyle (1993). Measuring resource damages with contingent valuation: tests of validity and reliability. In J. Hausman (Ed.), *Contingent Valuation: A Critical Assessment*, Contributions to economic analysis. Amsterdam, the Netherlands: Elsevier Science Publishers B.V.

Diamond, P. and J. Hausman (1993). On contingent valuation measurement of nonuse values. In J. Hausman (Ed.), *Contingent Valuation: A Critical Assessment*, Contributions to economic analysis. Amsterdam, the Netherlands: Elsevier Science Publishers B.V.

Frederick, S. and B. Fischhoff (1997, November). Scope (in)sensitivity in elicited valuations. Department of Social and Decision Sciences, Carnegie Mellon University. Working paper. Web: `http://stuff.mit.edu/people/shanefre/`.

Green, C. and S. Tunstall (1999). A psychological perspective. In I. Bateman and K. Willis (Eds.), *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*, pp. 183–206. Oxford, U.K.: Oxford University Press.

Hanemann, W. (1984). Discrete/continuous models of consumer demand. *Econometrica 52*(3), 541–562.

Hanemann, W. (1994). Valuing the environment through contingent valuation. *Journal of Economic Perspectives 8*(4), 19–43.

Hanemann, W. and B. Kanninen (1999). The statistical analysis of discrete-response CV data. In I. Bateman and K. Willis (Eds.), *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries.* New York, NY: Oxford University Press.

Herriges, J. and J. Shogren (1996). Starting point bias in dichotomous choice valuation with follow-up questioning. *Journal of Environmental Economics and Management 30*, 112–131.

Jorgensen, B., G. Syme, B. Bishop, and B. Nancarrow (1999). Protest responses in contingent valuation. *Environmental and Resource Economics 14*(1), 131–150.

Kahneman, D. (1986). Comments. In R. Cummings, D. Brookshire, and W. Schulze (Eds.), *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method.* Totowa, NJ: Rowman and Allanheld.

Kahneman, D. and J. Knetsch (1992). Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management 22*, 57–70.

Keith, J. and C. Fawson (1996). Compliance bias in dichotomous choice CVM: Some evidence from a Utah wilderness study. Department of Economics, Utah State University, Working Paper 9627. Web: `http://ideas.repec.org/p/wop/utstec/9627.html`.

Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy 74*(2), 132–157.

Li, C.-Z. and L. Mattsson (1995). Discrete choice under preference uncertainty: an improved structural model for contingent valuation. *Journal of Environmental Economics and Management 28*(2), 256–269.

Mendel, J. (2001). *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Upper Saddle River, NJ: Prentice Hall.

Messonnier, M., J. Bergstrom, C. Cornwell, R. Teasley, and H. Cordell (2000). Survey response-related biases in contingent valuation: concepts, remedies, and empirical application to valuing aquatic plant management. *American Journal of Agricultural Economics 83*(3), 438–450.

Moldoveanu, M. and E. Langer (2002). False memories of the future: a critique of the applications of probabilistic reasoning to the study of cognitive processes. *Psychological Review 109*(2), 358–375.

Nunes, P. and L. Onofri (2004, September). The economics of warm glow: A note on consumers behavior and public policy implications. Fondazione Eni Enrico Mattei, Sustainability Indicators and Environmental Valuation (SIEV) paper 113.2004. `Web: http://www.feem.it/Feem/Pub/Publications/WPapers/`.

Randall, A. and J. Hoehn (1996). Embedding in market demand systems. *Journal of Environmental Economics and Management 30*, 369–380.

Ready, R., S. Navrud, and W. Dubourg (2001). How do respondents with uncertain willingness to pay answer contingent valuation questions? *Land Economics 77*(3), 315–326.

Ready, R., J. Whitehead, and G. Blomquist (1995). Contingent valuation when respondents are ambivalent. *Journal of Environmental Economics and Management 29*(2), 181–196.

Rizzello, S. (1999). *The Economics of the Mind*. Edward Elgar Publishing.

Rubin, D. (1987). *Multiple imputation for nonresponse in surveys*. Wiley series in probability and mathematical statistics. New York, NY: Wiley.

Schkade, D. and J. Payne (1994). How people respond to contingent valuation questions: A verbal protocol analysis of willingness to pay for an environmental regulation. *Journal of Environmental Economics and Management 26*, 88–109.

Thurstone, L. (1927). A law of comparative judgment. *Psychological Review 34*, 273–286.

Train, K. (2003). *Discrete Choice Methods with Simulation*. New York, NY: Cambridge University Press.

Tversky, A. and D. Kahneman (1971). Belief in the law of small numbers. *Psychological Bulletin 76*(2), 105–110.

Van Kooten, G., E. Krcmar, and E. Bulte (2001). Preference uncertainty in non-market valuation: a fuzzy approach. *American Journal of Agricultural Economics 83*(3), 487–500.

Walker, J. (2001). *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables*. Ph. D. thesis, Massachusetts Institute of Technology.

Wang, H. (1997). Treatment of "don't know" responses in contingent valuation surveys: a random valuation model. *Journal of Environmental Economics and Management 32*(2), 219–232.

Wynand, P. and B. van Praag (1981). The demand for deductibles in private health insurance: A probit model with sample selection. *Journal of Econometrics 17*, 229–252.

Zadeh, L. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences 8*, 199–249.

Zadeh, L. (1999). Fuzzy logic = computing with words. In L. Zadeh and J. Kacprzyk (Eds.), *Computing with Words in Information/Intelligent Systems*, Volume 33 of *Studies in Fuzziness and Soft Computing*. New York, NY: Physica-Verlag.

Zeger, S., K. Liang, and P. Albert (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics 44*, 1049–1060.