WESTERN REGIONAL RESEARCH PUBLICATION

W-133
Benefits and Costs Transfer in Natural Resource Planning

Seventh Interim Report

June, 1994

Compiled by
Richard C. Ready

Department of Agricultural Economics
400 Agricultural Engineering Bldg.
University of Kentucky
Lexington, KY 40546

CONTENTS

INTRODUCTION

W-133, Benefits and Costs Transfer in Natural Resource Planning, is a Western Regional Research project chartered by United States Department of Agriculture, Cooperative States Research Service. Its official membership is comprised of researchers representing Land Grant Experiment Stations in 25 states from across the country. Members share a common research interest in developing methodology for non-market valuation and assessing validity of those methods, and in developing methodology for the transfer of non-market value estimates from studied sites to unstudied sites and assessing the validity of such benefits transfer. W-133 provides a framework within which members from different states can meet to plan and discuss cooperative research. In addition to official members, many other researchers from academia and from state and federal government agencies participate unofficially by attending and making presentations at meetings of the technical committee, and by working jointly with W-133 members on cooperative research projects. This interaction benefits members and non-members alike, and is one of the unique strengths of W-133.

The stated objectives of the W-133 project are: 1) to provide site specific use and non-use values of natural resources for public policy analysis, and 2) to develop protocols for transferring value estimates to unstudied sites. Ongoing research towards meeting these two objectives is targeted towards four resource policy areas: water based recreation, groundwater quality, wetlands, and recreational fisheries.

In February, 1994, W-133 held its annual technical committee meeting at the Westward Look Resort in Tucson, Arizona. Time was allotted for members to meet to plan and discuss progress on ongoing cooperative research. Additionally, several members and participating non-members presented papers based on individual and cooperative research that addresses the objectives of the project. This volume reproduces those papers into one accessible source.

The first set of papers in this volume deal with methodological and analytical issues surrounding valuation of outdoor recreation opportunities. The second set of papers addresses methodological and analytical issues involved in the contingent valuation method, including assessments of the validity of CVM. Case studies that describe innovations in methodology and analysis are included in both sets of papers. The third set of papers is from an invited papers session at the meetings, organized by John Loomis (Colorado State University), entitled "The Roles of Economics in Decisions About Endangered Species." These papers explore how economists can contribute to the process of decisionmaking regarding endangered species.

I would like to thank John Keith and David Plane from the Western Regional Science Association for their help in meeting arrangements, and thank Rita Parsons and Kristin Rehrman for their help in editing and formatting the papers included in this volume.

Richard C. Ready
University of Kentucky
June, 1994

TWO RUMs unCLOAKED:

Nested-Logit Models of Site Choice

and

Nested-Logit Models of Participation and Site Choice[1]

Edward R. Morey
Department of Economics
Campus Box 256
University of Colorado
Boulder, Colorado 80309-0256
Fax 303-449-3459

March 24, 1994

an earlier version was presented at the W-133 Meetings, Tucson Arizona, February, 1994

ABSTRACT

Nested logit is increasingly advocated as a tool of recreational demand and benefit estimation. The intent of this short monograph is to lay out, in a simple fashion, the theory behind the nested-logit model of site choice and the nested-logit model of participation and site choice; and then provide rigorous but straightforward derivations of the properties of nested-logit models including: the probability of choosing a particular alternative, likelihood functions, expected maximum utility, a compensating variation, and an equivalent variation. Also discussed are estimation, regularity conditions, the interpretation of the *scaling* parameters, and the relation between those scaling parameters and the *Independence of Irrelevant Alternatives - I.I.A.* assumption(s) imbedded in both the nested logit model and its special cases. While this primer does not derive any new theoretical results, it does provide a synthesis of materials that are widely diffused, sometimes misstated, and often in a form that is not readily accessible to econometricans working in the area of recreational demand and benefit estimation. Examples are used to tie the theory to recreational demand and benefit estimation.

[1]I want to thank George, Trudy, Cathy, Sally and Douglass for comments and advice on earlier versions.

Policy analysts often require the consumer's surplus (CV and/or EV) associated a change in the costs and/or characteristics of a group of consumption activities where the consumer's choice of consumption activity generally involves two simultaneous decisions; whether to participate in a given class of activities and, if so, which specific alternative to choose from that class. For example, one simultaneously decides both whether participate in a given class of site-specific recreational activities, and if so, which site to visit. Joint decisions of this type can be modelled in either a multinomial logit (MNL) framework or a nested-logit (NL) framework[2].

Use of the NL model, in contrast to the MNL model, is increasingly advocated; particularly when the intent is to simultaneously model both the decision to participate and the choice of site (see [3], [4], [6], [11], [21], [28] and [30]).[3] The argument is that the I.I.A. (Independence of Irrelevant Alternatives) assumption, implicit in the MNL model, while possibly *reasonable* when all the alternatives are recreational sites of a particular type, is not reasonable when the sites differ by type and/or one of the alternatives is nonparticipation. Participation and site choice should therefore be modelled as a two, or more, stage nested decision that does not impose I.I.A. a priori across all pairs of alternatives. For example, at stage one the individual decides whether to participate, and at stage two which site to visit if the individual chooses to participate.

The intent of this primer is not to derive any new results. Rather the intent is to lay out in a simple fashion the nested-logit model and then provide rigorous, but straightforward, derivations of its properties. Much of this material is widely diffused and often presented in such a general framework that it is not readily accessible to econometricans working in the area of recreational demand and benefit estimation. For example, McFadden [22] derives maximum expected utility from a generalization of the nested-logit model, but his proof is part of such a general argument that its usefulness in the derivation of consumer's surplus from nested-logit models of recreational

---

[2]Joint decisions of this type can also be modelled in other frameworks, but those other frameworks are not the topic of this paper.

[3]Additional examples of discrete-choice models of recreational demand between 1988 and the present are [5], [13], [15], [16], [24], [29], [31], [32] and [33]. Earlier examples are [12], [14], [17] and [25].

demand is not transparent. The derivations in McFadden [23] are also presented in very general terms. In addition, the literature is muddled by the discussion of the nested-logit model in the widely used reference book on discrete choice modelling by Ben-Akiva and Lerman [2]. Ben-Akiva and Lerman incorrectly assume that the nested-logit model can be derived from an error components model.[4] This incorrect assumption, that the nested-logit model can be derived from an error components model, is also utilized by me in an occasionally cited discussion paper [26]. Given all this diffusion, a short, but rigorous, monograph on the properties of nested-logit seems in order

Section I derives the probability of choosing an alternative, and then uses it to form some sample-specific likelihood functions. Section II interprets the parameters in the nested logit as they relate to unobserved attributes and the dependence, or independence, of the random components of utility. In this framework, the I.I.A. assumption is discussed. Section II also identifies and discusses special cases of nested-logit. Section III advocates *Full Information Maximum Likelihood - FIML* estimation and discourages two-step sequential estimation. Section IV derives expected maximum utility, Section V discusses budget exhaustion and the conditions it imposes on the conditional indirect utility functions for the alternatives, and Section VI uses expected maximum utility to derive a compensating variation and an equivalent variation. Section VII expands the nest to three levels. Example boxes are used throughout to tie the theory to the application of recreation demand and benefit estimation.

[4]Specifically they assume (page 287) that the random term in the conditional indirect utility function for alternative mj can be divided into two components where one of the components has an Extreme Value distribution, and the sum of the random components has an Extreme Value distribution. However, there is no proof that there is any distribution for the second random term that would cause the sum of the two components to have an Extreme Value distribution, and in general one cannot expect there to be such a distribution.

I. The Two-Level Nested-Logit Model of Recreational Demand: Its CDF, Probabilities, and Likelihood Function

The intent of this section is to use the basics of probability theory to derive the probability of choosing each alternative from the assumptions that form the basis of the two-level nested logit model of consumer demand. Once accomplished these probability equations can be used to form likelihood functions, the specific form of the likelihood function depending on the properties of one's sample.

> **Examples:** Consider two different two-dimensional models of recreational demand: a model of participation and site choice; and a model of just site choice, but where the sites are of three distinct types.
>
> **A model of participation and site choice:** Consider a choice set with C=10 alternatives, staying at home, going bowling, and visiting one of eight fishing sites. In which case, one might assume M has two elements; 1= fishing and 2=not fishing; J_1=8 (the number of sites) and J_2=2 (staying home and bowling).
>
> **A Model of Site Choice with Saltwater Sites, Lakes and Rivers:**
> Consider a choice set with C=12 alternatives, three saltwater sites, four lakes and five rivers. In which case, one might assume M has three elements; 1= saltwater fishing, 2=lake fishing; and 3=river fishing; J_1=3, J_2=4 and J_3=5.

The two-level nested-logit model is designed to explain an individual's choice of alternative when there is a two-dimensional choice set from which the individual **must** choose from one of C distinct alternatives; where one of the dimensions of the choice set can be characterized in terms of M distinct types, and the other dimension J distinct types, C $\leq$ M×J. The individual chooses an alternative, ni, where n $\in$ M and i $\in$ J, subject to the restriction that their choice of type in terms of the J dimension has to be consistent with their choice of type in terms of the M dimension. Without loss of generality, nest the two dimensions such that if the individual chooses an alternative of type n $\in$ M, then the individual's choice of alternative in the J dimension is restricted to a subset of the J types where this subset has J_n elements, where J_m is the number of J types consistent with a choice of type m $\in$ M.

Nested-logit models assume the utility the individual receives if he chooses alternative mj is

$$(1) \quad U_{mj} = V_{mj} + \epsilon_{mj} \qquad \forall\ (mj) \in C$$

where V_{mj} is the systematic component of utility and ϵ_{mj} is a random component. Both terms are known to the individual but the ϵ_{mj} are unobserved by the researcher so are random variables from the researcher's perspective. Let $<\epsilon_{mj}>$ denote the vector of these C random terms; that is $<\epsilon_{mj}> = \{\epsilon_{11},\epsilon_{12},...,\epsilon_{1J_1},\epsilon_{21},\epsilon_{22},...,\epsilon_{2J_2},...,\epsilon_{M1},\epsilon_{M2},...,\epsilon_{MJ_M}\}$. Let $f(<\epsilon_{mj}>)$ denote their joint density function (DF),and let $F(<\epsilon_{mj}>)$ denote their cumulative density function (CDF).[5]

In which case, the probability of choosing a particular alternative is derived by noting that

(2) $Prob(ni) = Prob[U_{ni} > U_{mj} \ \forall \ mj \neq ni] = Prob[\epsilon_{mj} < V_{ni} - V_{mj} + \epsilon_{ni} \ \forall \ mj \neq ni]$

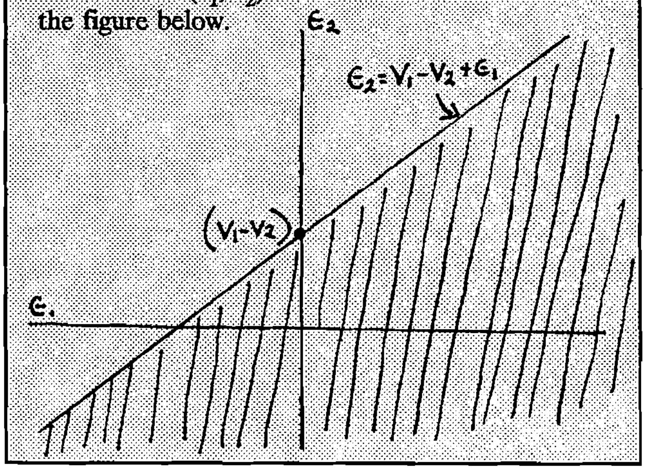
Without loss of generality, order the alternatives so that alternative *ni* is the first alternative; i.e., 11. Therefore.

(3) $Prob(11) = Prob[\epsilon_{mj} < V_{11} - V_{mj} + \epsilon_{11} \ \forall \ mj \neq 11]$

$$= \int_{\epsilon_{11}=-\infty}^{+\infty} \int_{\epsilon_{12}=-\infty}^{V_{11}-V_{12}+\epsilon_{11}} \cdots \int_{\epsilon_{mj}=-\infty}^{V_{11}-V_{mj}+\epsilon_{11}} \cdots \int_{\epsilon_{MJ}=-\infty}^{V_{11}-V_{MJ}+\epsilon_{11}} f(\epsilon_{11},\epsilon_{12},...,\epsilon_{mj},...,\epsilon_{MJ})\, d\epsilon_{JM} ... d\epsilon_{mj} ... d\epsilon_{12}\, d\epsilon_{11}$$

This is just the area under the density function, $f(<\epsilon_{mj}>)$, where $U_{11} > U_{mj} \ \forall \ mj \neq 11$. While Equation (3) is an straightforward representation of Prob(11), Prob(11) can be represented more compactly in terms of the CDF, $F(<\epsilon_{mj}>)$. Equation (3) expresses Prob(11) as a C-level multiple integral; using the CDF, Prob(11) can alternatively be expressed as a single integral. The

This area under $f(<\epsilon_{mj}>)$ can be visualized when C has only two elements (e.g., a trip to site 1 or a trip to site two), $f(<\epsilon_{mj}>) = f(\epsilon_1,\epsilon_2)$ and the probability of visiting site 1, Prob(1), is the area under $f(\epsilon_1, \epsilon_2)$ above the shaded area in the figure below.



[5]Two things should be noted about equation (1). Equation (1) is restrictive in that one could hypothesize two-level discrete choice models that do not fulfill equation (1) (e.g., $U_{mj} = V_{mj}\epsilon_{mj}$). Secondly, there are discrete-choice models that fulfill equation (1) that are not nested-logit models; e.g., multivariate Probit models.

ability to express Prob(11), and more generally Prob(ni), as a single integral makes evaluation of these probability functions much more tractable.[6]

The first step in expressing Prob(11) in terms of the CDF is to note that, in general,

(4) $Prob[\epsilon_{mj} < \overline{\epsilon_{mj}} \ \forall mj,\ m=1,...M.,\ j=2,...J.\ :\ \epsilon_{11} = \overline{\epsilon_{11}}]$

$$= \int_{\epsilon_{12}=-\infty}^{\overline{\epsilon_{12}}} \int_{\epsilon_{mj}=-\infty}^{\overline{\epsilon_{mj}}} \cdots \int_{\epsilon_{MJ}=-\infty}^{\overline{\epsilon_{MJ}}} f(\overline{\epsilon_{11}}, \epsilon_{12}, ..., \epsilon_{mj}, ..., \epsilon_{MJ}) d\epsilon_{JM} ... d\epsilon_{mj} ... d\epsilon_{12}$$

$$= F_{11}(\overline{\epsilon_{11}}, \overline{\epsilon_{12}}, ..., \overline{\epsilon_{mj}}, ..., \overline{\epsilon_{MJ}})$$

where $F_{ni}(.)$ denotes the derivative of F with respect to its $(ni)^{th}$ argument, and the *bar over a variable*, ¯, just denotes a specific value of that variable. Equation 4 tells us that the area under the density function defined in the middle term of equation (4), which is a probability, can be expressed as a derivative of the CDF. The probability that $[\epsilon_{mj} < \overline{\epsilon_{mj}} \ \forall mj,\ m=1,...M.,\ j=2,...J.]$, is then obtained by integrating equation (4) with respect to ϵ_{11} from minus to plus infinity; that is

(5) $Prob[\epsilon_{mj} < \overline{\epsilon_{mj}} \ \forall mj,\ m=1,...M.,\ j=2,...J.] = \int_{\epsilon_{11}=-\infty}^{+\infty} F_{11}(\epsilon_{11}, \overline{\epsilon_{12}}, ..., \overline{\epsilon_{mj}}, ..., \overline{\epsilon_{MJ}}) d\epsilon_{11}$

Utilizing Equations (3) and (5), the probability of choosing alternative 11 is, in terms of the CDF,

(6) $Prob(11) = Prob[\epsilon_{mj} < V_{11} - V_{mj} + \epsilon_{11} \ \forall mj,\ m=1,...M.,\ j=2,...J.]$

$$= \int_{\epsilon_{11}=-\infty}^{+\infty} F_{11}(<V_{11}-V_{mj}+\epsilon_{11}>) d\epsilon_{11}$$

where $<V_{11}-V_{mj}+\epsilon_{11}> = \{\epsilon_{11}, V_{11}-V_{12}+\epsilon_{11}, ..., V_{11}-V_{mj}+\epsilon_{11}, ..., V_{11}-V_{MJ}+\epsilon_{11}\}$. However, since there is nothing unique about alternative 11,

(7) $$Prob(ni) = \int_{\epsilon_{ni}=-\infty}^{+\infty} F_{ni}(<V_{ni}+\epsilon_{ni}-V_{mj}>) d\epsilon_{ni}$$

where $<V_{ni}-V_{mj}+\epsilon_{ni}> = \{V_{ni}-V_{11}+\epsilon_{ni}, V_{ni}-V_{12}+\epsilon_{ni}, ..., V_{ni}-V_{ni}+\epsilon_{ni}, ..., V_{ni}-V_{MJ}+\epsilon_{ni}\}$.

---

[6]Who likes to evaluate multiple integrals?

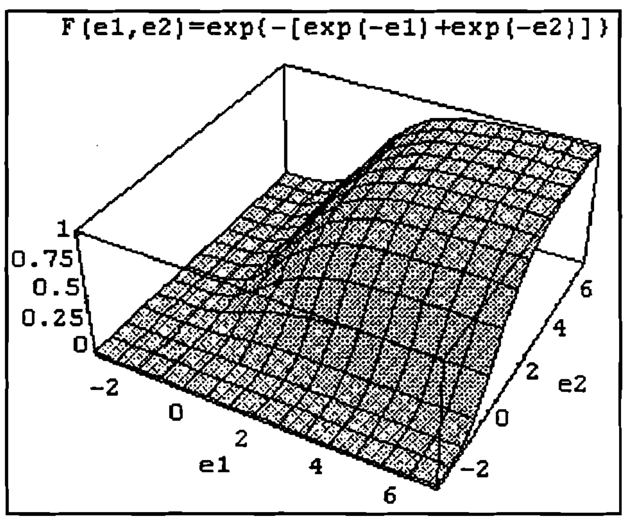
As noted above, the joy of Equation (7) over Equation (3) is Equation (7) is a single integral, whereas Equation (3) is a C-level multiple integral. Equation (7) is the probability of choosing alternative ni for any model that assumes equation (1). Up to this point the model is very general; it is consistent with any $F(<\epsilon_{mj}>)$.

To generate a two-level nested-logit model, specifically assume that the CDF is

$$(8)^{7} \quad F(<\epsilon_{mj}>) = \exp\{-\sum_{m=1}^{M} a_m [\sum_{j=1}^{J_m} e^{-s_m \epsilon_{mj}}]^{(1/s_m)}\}$$

where $a_m > 0$ and $s_m \geq 1 \; \forall \; m$.[8] This CDF is a special case of a Multivariate Generalized Extreme Value Distribution. The Generalized Extreme Value Distribution was first proposed by McFadden [22]. The task at hand is to show that the derivative of the this CDF when plugged into equation (7) generates the Prob(ni) equation for the two-level nested logit model.

A major reason for choosing this particular CDF, Equation (8), is that when Equation (8) is assumed, Equation (7), Prob(ni), has a closed-form solution. This greatly simplifies estimation of the model, eliminating the need for numerical integration. Most cumulative density functions do not generate closed forms for the Prob(ni). For example, if one assumes



---

[7]Alternatively, if one assumed a multivariate normal CDF, the model would be multivariate Probit.

[8]The restriction that $a_m > 0$ and $s_m \geq 1 \; \forall \; m$ is sufficient, but not necessary, to imply that Equation (8) is a well-behaved CDF; that is, to imply Equation (8) never takes on negative values, is monotonic in its arguments, and its value never exceeds one. The implications of violating this restrictions is discussed in Section III. Those familiar with McFadden [22] will note that I have broken with tradition and not used his notation. My s_m is his $1/(1 - \sigma_m)$. I find my notation simpler both in terms of word-processing and comprehension. Note that $s_m \geq 1 \Leftrightarrow 0 \leq \sigma_m < 1$.

the CDF is multivariate normal (the multivariate Probit model), the Equation (7) integral will not have a closed- form solution, so estimation of the likelihood function requires complex, numerical, multiple integration. This is why estimated multivariate probit models limit the choice set to a small number of alternatives (two, three, four), but nested-logit models can be estimated with large numbers of alternatives.

Examples: Consider, the $F(<\epsilon_{mj}>)$ for the two models of recreational demand introduced in the first example box.

The model of participation and site choice where M has two elements; 1= fishing and 2=not fishing; $J_1=8$ (the number of sites) and $J_2=2$ (staying home and bowling). For this example,

$$F(<\epsilon_{mj}>) = \exp\{-(a_1[\exp(-s_1\epsilon_{11})+\ldots+\exp(-s_1\epsilon_{18})]^{(1/s1)} + a_2[\exp(-s_2\epsilon_{21})+\exp(-s_2\epsilon_{22})]^{(1/s2)})\}$$

The Model of Site Choice with Saltwater Sites, Lakes and Rivers, where M has three elements; 1= saltwater fishing, 2=lake fishing; and 3=river fishing; $J_1=3$, $J_2=4$ and $J_3=5$. For this example,

$$F(<\epsilon_{mj}>) = \exp\{-(a_1[\exp(-s_1\epsilon_{11})+\ldots+\exp(-s_1\epsilon_{13})]^{(1/s1)} + a_2[\exp(-s_2\epsilon_{21})+\ldots+\exp(-s_2\epsilon_{24})]^{(1/s2)} + a_3[\exp(-s_3\epsilon_{31})+\ldots+\exp(-s_3\epsilon_{35})]^{(1/s3)})\}$$

To obtain the closed form of the Prob(ni) equation, first take the derivative of the Multivariate Extreme Value CDF with respect to its $(ni)^{th}$ element. One obtains

$$(9)\quad F_{ni}(<\epsilon_{mj}>) = \exp\{-\sum_{m-1}^{M} a_m[\sum_{j-1}^{J_m} e^{-s_m(\epsilon_{ml})}]^{(1/s_m)}\} \; a_n[\sum_{j-1}^{J_n} e^{-s_n\epsilon_{nj}}]^{(1/s_n)-1} \; e^{-s_n\epsilon_{ni}}$$

Substituting $<V_{ni}+\epsilon_{ni}-V_{mj}>$ for $<\epsilon_{mj}>$ in equation (9), one obtains

$$(10)\quad F_{ni}(<V_{ni}+\epsilon_{ni}-V_{mj}>)\ = \exp\{-\sum_{m-1}^{M} a_m[\sum_{j-1}^{J_m} e^{-s_m(V_{ni}+\epsilon_{ni}-V_{mj})}]^{(1/s_m)}\}$$

$$\{a_n[\sum_{j-1}^{J_n} e^{-s_n(V_{ni}-V_{nj}+\epsilon_{ni})}]^{(1/s_n)-1} e^{-s_n\epsilon_{ni}}\}$$

Before substituting the RHS of equation (10) into equation (7) to obtain Prob(ni), simplify equation (10) into terms that do, and don't, involve ϵ_{ni} so that $F_{ni}(<V_{ni}+\epsilon_{ni}-V_{mj}>)$ in equation (7) will be easy to integrate with respect to ϵ_{ni}. Factoring equation (10) one obtains

$$(11)\quad F_{ni}(<V_{ni}+\epsilon_{ni}-V_{mj}>)\ = e^{-\epsilon_{ni}} \exp\{-e^{-\epsilon_{ni}} e^{-V_{ni}} B\}\ A$$

where

$$(12)\quad A\ = a_n[\sum_{j=1}^{J_n} e^{s_n V_{nj}}]^{(1/s_n)-1}\ e^{-V_{ni}}\ e^{s_n V_{ni}} \qquad \text{and}$$

$$(13)\quad B\ = \sum_{m=1}^{M} a_m[\sum_{j=1}^{J_m} e^{s_m V_{mj}}]^{1/s_m}$$

Note that A and B do not depend on ϵ_{ni} . Plugging equation (11) into equation (7) one obtains

$$(14)\quad Prob(ni)\ = A\ \int_{\epsilon=-\infty}^{+\infty} e^{-\epsilon}\ \exp\{-e^{-\epsilon}\ F\ B\}\ d\epsilon$$

where $F\ = e^{-V_{ni}}$. Rather than trying to integrate this with respect to ϵ, simplify it further by making the change of variables $m\ = e^{-\epsilon}$ ($\Rightarrow d\epsilon\ = -(1/m)dm$) to obtain

$$(15)\quad \boldsymbol{Prob(ni)}\ = \boldsymbol{A} \int_{m=0}^{\infty} \exp\{-\boldsymbol{mFB}\}\ \boldsymbol{dm}\ = \boldsymbol{A/(FB)}$$

Substituting back in for A, B, and F, one obtains

$$(16)\quad Prob(ni) = \frac{e^{s_n V_{ni}}\, a_n[\sum_{j=1}^{J_n} e^{s_n V_{nj}}]^{(1/s_n)-1}}{\sum_{m=1}^{M} a_m[\sum_{j=1}^{J_m} e^{s_m V_{mj}}]^{1/s_m}}$$

which is the probability of choosing alternative ni in a two-level nested logit model. Inclusion of the $<a_{mj}>$ parameters is equivalent to adding a group-specific constant term, α_m, to each of the V_{mj}, where $\alpha_m = \ln(a_m)$. To see this replace, replace a_m with e^{α_m} ($\rightarrow \alpha_m = \ln(a_m)$); in which case Prob(ni) can be rewritten as

$$(16a)\quad Prob(ni) = \frac{e^{s_n(\alpha_n + V_{ni})}\, [\sum_{j=1}^{J_n} e^{s_n(\alpha_n + V_{nj})}]^{(1/s_n)-1}}{\sum_{m=1}^{M} [\sum_{j=1}^{J_m} e^{s_m(\alpha_m + V_{mj})}]^{1/s_m}}$$

Example: For the **Model of Site Choice with Saltwater Sites, Lakes and Rivers**, the probability that the individual will choose the third lake site (ni = 23) is
$Prob(23) =$
$\exp(s_2 V_{23}) a_2[\exp(s_2 V_{21}) + ... + \exp(s_2 V_{24})]^{((1/s2)-1)}$
divided by
$\{a_1 [\exp(s_1 V_{11}) + ... + \exp(s_1 V_{13})]^{1/s1}$
$+ a_2[\exp(s_2 V_{21}) + ... + \exp(s_2 V_{24})]^{1/s2}$
$+ a_3[\exp(s_3 V_{31}) + ... + \exp(s_3 V_{35})]^{1/s3}\}$

If desired, the probability, equation (16), can be decomposed into the probability of choosing an alternative of type n multiplied by the probability of choosing alternative i from the group of alternatives that are of type n; i.e.,

$$(17)\quad Prob(ni) = Prob(i|n)Prob(n)$$ where

$$(18)\quad Prob(n) = \frac{a_n[\sum_{j=1}^{J_n} e^{s_n V_{nj}}]^{1/s_n}}{\sum_{m=1}^{M} a_m[\sum_{j=1}^{J_m} e^{s_m V_{mj}}]^{1/s_m}}$$ and

$$(19)\quad Prob(i|n) = \frac{e^{V_{ni}}}{[\sum_{j=1}^{J_n} e^{V_{nj}}]}$$

Equation (16) is made explicit by specifying functional forms form for the V_{mj}, where V_{mj} is the conditional indirect utility function for alternative mj. V_{mj} is typically assumed some function, often linear, of the cost of alternative mj, the budget, and the characteristics of alternative mj. If, for example, mj is a fishing site, the variables might be the cost of a trip to site mj, expected catch rate at site mj and other characteristics of the site. The regularity conditions on the V_{mj} are considered in Section V.

Consider now the problem of estimating the parameters in the V_{mj} functions using a sample of individuals that reports the alternative, or alternatives, chosen by each individual in the sample. At this point it is important to make a distinction. Denote each time an individual must choose between the C alternatives in the choice set a *choice occasion*. An important distinction is whether the sample contains information on the alternative chosen for just one choice occasion for each individual, or whether the data set reports, by individual, the alternative chosen on each of a number of choice occasions. The number of observed choice occasions could vary across individuals. Start with the simpler case where the sample contains the choice on only one choice occasion for each individual. The

> **Examples:** Consider what constitutes a choice occasion for the two models of recreation demand introduced in the first example box.
>
> In the model of **participation and site choice,** the fishing season can be divided into a finite number of periods and each period is a choice occasion. For example, each week in the season might be defined as a choice occasion, or, more generally, the season might be divided into a fixed number of periods, but with no restriction that each period is of a specified length. The critical issue for this simple model of participation and site choice is that no more than one site is chosen on each choice occasion.
>
> In the model of **site choice with saltwater sites, lakes and rivers,** a choice occasion is each time a fishing trip is taken.

probability of observing individual h choosing alternative ni on the one choice occasion is Prob(hni). If one further assumes that the choices of the H individuals in the sample are statistically independent; that is, $cov(\epsilon_{hni}, \epsilon_{kmj}) = 0, \ \forall$ n,i,m,j, h≠k, the probability of observing the set of observed choices is determined by the multinomial density function

$$(20) \quad f(<y_{1mj}>,<y_{2mj}>,...,<y_{Hmj}>) = \frac{H!}{HC} \prod_{h=1}^{H} \prod_{n=1}^{M} \prod_{i=1}^{J_n} Prob(hni)^{y_{hni}}$$

where , $y_{hni} = 1$ if individual *h* choose alternative *ni*, and zero otherwise, C is the number of alternatives in the choice set, Prob(hni) is defined by equation (16), and

$$<y_{hmj}> = \{y_{h11},y_{h12},...,y_{h1J_1},y_{h21},y_{h22},...,y_{h2J_2},...,y_{hM1},y_{hM2},...,y_{hMJ_M}\}.$$

Equation (20) is the likelihood function for this sample; that is, it is the probability of observing the choices in the sample as a function of the Prob(hmj). The task is to find those values of the parameters in the Prob(hmj) that maximize the likelihood function. Since the parameters that maximize the log of the likelihood function also maximize the likelihood function, estimation is simplified by finding those values of the parameters that maximize the log of the likelihood function.[9]

$$(21) \quad L = \sum_{h=1}^{H} \sum_{n=1}^{M} \sum_{i=1}^{J_n} y_{hni} \ln[Prob(hni)]$$

Estimation of log likelihood functions are briefly discussed in section III.

Consider now the case where the data set reports, by individual, the alternative chosen on each of a number choice occasions, where the number of observed choice occasions may vary across individuals. Such samples are generated by *repeated-choice* problems; that is the discrete-choice problem faced by the each individual repeats, so there are multiple choice occasions. Some discrete

[9]The additive term ln[H!/(HC)] is omitted because it does not depend on the values of the parameters in the Prob(hmj).

choice problems such as what furnace to purchase or what individual to marry do not repeat, or, hopefully, do not repeat to often. In contrast, discrete choice problems in recreational demand are characterized by repetition. The problem of where to go on a fishing trip repeats every time one takes a trip. The problem of whether to take a fishing trip also repeats every choice occasion (for example, every day or every week).

Examples: The log likelihood functions for the two models of recreational demand introduced in the first example box when the sample only reports the alternative chosen on one choice occasion for each individual in a sample of 100 individuals.

The model of participation and site choice where M has two elements; 1= fishing and 2=not fishing; $J_1=8$ (the number of sites) and $J_2=2$ (staying home and bowling). For this example,

$L = \Sigma_h \Sigma_n \Sigma_j \; y_{hni} \ln(\text{Prob}(hni)) \}$

where h=1 to 100; n = 1,2;
$J_1 = 8$ so $j_1 = 1$ to 8; and $J_2 = 2$ so $j_2 = 1,2$.

The Model of Site Choice with Saltwater Sites, Lakes and Rivers, where M has three elements; 1= saltwater fishing, 2=lake fishing; and 3=river fishing; $J_1=3$, $J_2=4$ and $J_3=5$. For this example,

$L = \Sigma_h \Sigma_n \Sigma_j \; y_{hni} \ln(\text{Prob}(hni)) \}$

where h=1 to 100; n = 1,2,3;
$J_1 = 3$ so $j_1 = 1$ to 3; $J_2 = 4$ so $j_2 = 1$ to 4; and $J_3 = 5$ so $j_3 = 1$ to 5.

Let T_h denote the number of choice occasions observed for individual h and ϵ_{hmjt} the random component in the utility individual h receives during choice occassion t if alternative mj is chosen. Assume, in addition to the previous assumption that choices are statistically independent across individuals, that choices, for a given individual, are statistically independent across choice occasions. That is, assume $\text{cov}(\epsilon_{knis}, \epsilon_{kmjt}) = 0, \; \forall$ k,n,i,m,j, s≠t.

For this case, let $Y_{hni} \equiv$ the number of time individual h chooses alternative ni, where $\sum_{n=1}^{M}\sum_{i=1}^{J_n} Y_{hni}=T_h$. The probability of observing the vector of alternatives$<Y_{hni}>$ for individual h is

determined by the multinomial density function.

$$(22)\quad f(<Y_{hmj}>) = \frac{T_h!}{\prod_{n=1}^{M}\prod_{l=1}^{J_n} Y_{hnl}!} \prod_{n=1}^{M}\prod_{l=1}^{J_n} Prob(hnl)^{Y_{hnl}}$$

The log of the likelihood function for this sample is

$$(23)\quad L = \sum_{h=1}^{H}\sum_{n=1}^{N}\sum_{l=1}^{J_n} Y_{hnl}\ \ln[Prob(hnl)]$$

Another common but more complicated type of sample is a sample that contains, for each individual, information on the specific alternative chosen for some choice occasions, but only partial information on the alternative chosen for other choice occasions. For example, one might know the specific alternative chosen for some choice occasions but for others only know which of the M groups the alternative is in. The log of the likelihood function for such an "incomplete" sample is

> **For example,** in our **model of participation and site choice** one might know for some choice occasions both whether and where an individual fished, but for other choice occasions only know that a trip was taken but have not information about the destination.
>
> In the **model of site choice for saltwater sites, lakes and rivers,** one might know for some choice occasions the exact site chosen, but for other choice occasions only know that the trip was to a river.

$$(24)\quad L = \sum_{h=1}^{H}\sum_{n=1}^{M}\{Y_{hn}\ln[Prob(hn)] + \sum_{l=1}^{J_n} Y_{hnl}\ \ln[Prob(hnl)]\}$$

where Prob(hn) is defined in equation (18), Y_{hn} is the number of times individual i is known to choose an alternative of type n where it is not known which of the J_n alternatives in group n was chosen. Note that in this case $T_h = \sum_{n=1}^{N}\sum_{l=1}^{J_n} Y_{hnl} + \sum_{n=1}^{N} Y_{hn}$.

II. Special Cases of the Two-Level Nested

The significance of the a_m and s_m parameters in the CDF, equation (8), are deciphered by remembering that the utility an individual receives if they choose alternative *ni* is, from the analyst's perspective, a random variable; i.e., $U_{ni} = V_{ni} + \alpha_m + \epsilon_{ni}$ where $\boldsymbol{\alpha_m} = \ln(a_m)$, $(V_{ni} + \alpha_m)$ is deterministic and ϵ_{ni} is the random variable. The V_{ni} are a function of *attributes* of the alternatives that are observed by the analyst. The $(\alpha_n + \epsilon_{ni})$ result from the impacts of the attributes that are not observed. As noted earlier, inclusion of the $<a_m>$ parameters is equivalent to adding a group-specific constant term, α_m, to each of the V_{mj}, where $\boldsymbol{\alpha_m} = \ln(a_m)$; this will be elaborated on below.

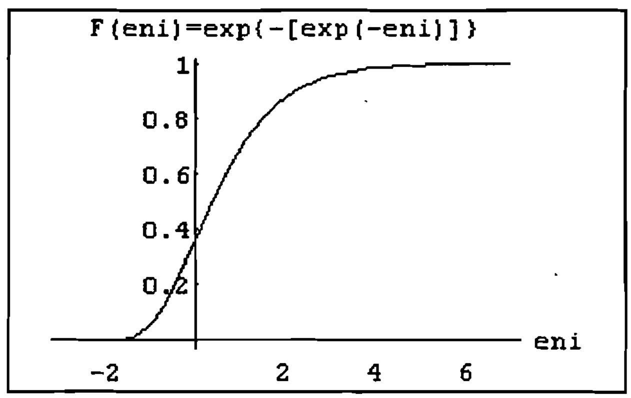
A critical issue in all two-level discrete choice models is whether each element of the vector $<\epsilon_{mj}>$ is **independently** drawn from the same univariate distribution, or whether elements of $<\epsilon_{mj}>$ are drawn from a multivariate distribution and therefore correlated. The nested-logit CDF, equation (8) allows the ϵ_{mj} to be correlated by type.

For example, in our model of participation and site choice, one would expect the random terms in the conditional indirect utility function for the fishing sites ($<\epsilon_{1i}>$) to be more correlated with one another than they are with the random term for staying at home (ϵ_{21}) or the random term for bowling (ϵ_{22}). This is because there are important, unobserved attributes of the alternatives that vary more, or less, across the fishing sites than they vary across the fishing sites and the other two alternatives. For example, the attribute fish stock varies across fishing sites but is always zero for staying at home and bowling, so omitting it would cause the ($<\epsilon_{1i}>$) to be more correlated with one another than they are with the (ϵ_{21}) or (ϵ_{22}).

Consider our **model of site choice with saltwater sites, lakes and rivers.** Assume, that aquatic vegetation varies significantly across rivers but not as much across lakes sites or saltwater sites. If this attribute of sites is unobserved, the random terms for the river alternatives ($<\epsilon_{3i}>$) will be more correlated with each other than are correlated with the random terms for lakes and saltwater sites.

What would cause the random terms, $<\epsilon_{mj}>$, to be correlated by type? If an attribute that is an important determinant of choice is not observed, it influences the magnitude of the $<\epsilon_{mj}>$, the $<\alpha_m>$, or both. If this attribute **varies** across alternatives within a group less, or more, than across alternatives in different groups, the random elements in group n will be more correlated with each other than they are with the random elements for alternatives that are not in group n.[10] If, in addition, the amount the unobserved attribute **varies** within a group varies by group, alternatives in some groups will be more correlated with each other than the alternatives in other groups are correlated with each other. In these two cases, it is inappropriate to assume that the random terms for all C alternatives are independently drawn from the same univariate distribution. The nested-logit CDF, equation (8), allows the random terms to be correlated by groups and for the degree of correlation to vary by group.

Alternatively, if the **variation** in the unobserved attributes is not systematic by group type, it is reasonable to assume that each element of $<\epsilon_{mj}>$ is drawn from the same univariate distribution. This is what is assumed by the multinomial logit model. It assumes that each element of $<\epsilon_{mj}>$ is independently drawn from a univariate Extreme Value Distribution with scale parameter κ.



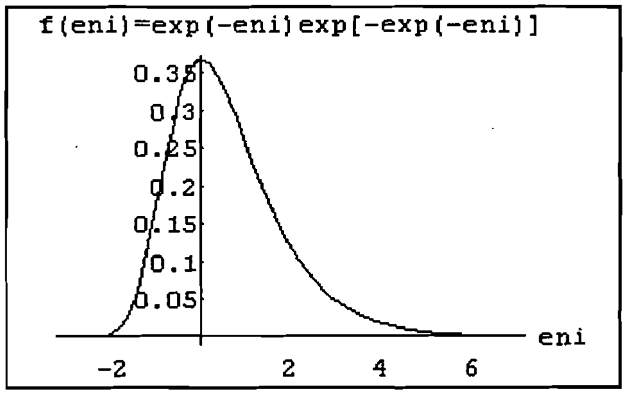
(25) $F(\epsilon_{ni}) = \exp\{-\kappa e^{-\epsilon_{ni}}\}$

The multinomial logit model is a special case of the two-level nested logit model and may be derived from equation (8) by restrictively assuming that a_m = a and s_m = 1 $\forall$ m. In which case, the probability of choosing alternative *ni,* equation (16) simplifies to

(26) $$Prob(ni) = \frac{e^{V_{ni}}}{\sum_{m=1}^{M}\sum_{j=1}^{J_m} e^{V_{mj}}}$$

[10]For example, if size is an important, unobserved attribute of the alternatives and size varies less within groups than across groups, the random terms for the alternatives that belong to type n will be more correlated with each other than they are with the random terms of alternatives that are not of type n

As is well known, the multinomial logit model imposes the I.I.AA assumption which says that the ratio of any two probabilities is independent of any change in any third alternative; that is



(27) $$\frac{Prob(ni)}{Prob(lk)} = \frac{e^{V_{ni}}}{e^{V_{lk}}}$$

This restriction is correct if the variation in unobserved attributes is not systematic by group type, but inappropriate if it is.

If an important unobserved attribute has the same magnitude for all the alternatives in a group but differs across groups, this will affect the $<a_{mj}>$ but not the $<s_{mj}>$. Such attributes cause alternatives within a group to be more similar to one another than they are to alternatives in different groups, but does not influence the correlations of the $<a_{mj}>$. Consider the following example. Size varies across alternative, but all alternatives of the same type are the same size. This omitted factor would cause the $<a_m>$ to vary in magnitude, but would not cause the elements of $<s_m>$ to differ from one.

The a_m and s_m add systematic variation across groups that is in addition to the systematic variation in terms of the V_{mj}; that is, a_m and s_m allow the groups to differ in systematic ways in addition to the differences that can be attributed to variations in the observed attributes that appear as independent variables in the V_{mj}. a_n reflects the relative attractiveness of alternatives of type *n*. Ceteris paribus, a_n will be large, in a relative sense, if alternatives of type *n* have more of an important, but unobserved, attribute.

Note that allowing a_m to vary, $s_m = 1 \; \forall \; M$, is not be sufficient to weaken the I.I.A assumption. This can be seen be considering a case where $s_m = 1 \; \forall \; M$ but a_m varies. In this case,

$$(28) \quad Prob(ni) = \frac{a_n e^{V_{ni}}}{\sum_{m=1}^{M} a_m \sum_{j=1}^{J_m} e^{V_{mj}}} \quad \text{and} \quad \frac{Prob(ni)}{Prob(lk)} = \frac{a_n e^{V_{ni}}}{a_l e^{V_{lk}}}$$

Thus, the I.I.A. assumption still holds for all pairs of alternatives given that a_n and a_l are treated as parameters. I.I.A. remains because the $<a_m>$ do not cause the elements of $<\epsilon_{mj}>$ to be correlated.

The s_n parameter, not a_n, is what is picking up part of the common (correlated) component in the random terms for all the alternatives of type *n*. The $<s_m>$ determine the extent to which the I.I.A. assumption is imposed. As noted above, $s_m = 1 \ \forall \ M$ imposes I.I.A across all pairs of alternatives. Alternatively, if $s_m \neq 1 \ \forall \ M$, the I.I.A. assumption will not be imposed across all pairs of alternatives, just some pairs.

Consider the two-level nested logit model with $a_m = a$ and $s_m = s \neq 1 \ \forall \ M$. In which case Prob(ni), equation (16), simplifies to

$$(29) \quad Prob(ni) = \frac{e^{sV_{ni}} \left[\sum_{j=1}^{J_n} e^{sV_{nj}}\right]^{(1/s)-1}}{\sum_{m=1}^{M} \left[\sum_{j=1}^{J_m} e^{sV_{mj}}\right]^{1/s}}$$

If $s_m = s \neq 1 \ \forall$ m, the random terms in each group are more correlated with each other than they are with the random terms in other groups, but the degree to which they are more correlated with their fellow group members is constant across groups. In nested logit models of recreational demand it is common to assume $s_m = s \neq 1 \ \forall$ m.

When $a_m = a$ and $s_m = s \neq 1 \ \forall$ M

$$(30) \quad \frac{Prob(ni)}{Prob(lk)} = \frac{e^{sV_{ni}}[\sum_{j=1}^{J_n} e^{sV_{nj}}]^{(1/s)-1}}{e^{sV_{lk}}[\sum_{j=1}^{J_l} e^{sV_{lj}}]^{(1/s)-1}}$$

Examining equation (30), one sees that

1. I.I.A. still holds for any pair of alternatives within the same group ($n = l$). If $n=l$, equation (30) simplifies to $\frac{Prob(ni)}{Prob(nk)} = \frac{e^{V_{ni}}}{e^{V_{nk}}}$.

2. I.I.A. still holds for all pairs of alternatives that are in different groups ($n \neq l$) if the alternative changed is not in the same group as either alternative in the pair. That is, equation (30) is not a function of changes in alternatives that are in other groups.

but,

3. I.I.A. does not hold for pairs of alternatives that are in different groups ($n \neq l$) if the alternative changed is in the same group as one of the alternatives in the pair. That is, equation (30) is a function of the attributes of alternatives n and l, so a change in any alternative in either group n or l will affect the ratio.

For example, if alternative 11 is altered,

it will not change $\frac{Prob(12)}{Prob(13)}$, $\frac{Prob(21)}{Prob(22)}$ or $\frac{Prob(31)}{Prob(42)}$, but will effect $\frac{Prob(12)}{Prob(22)}$. In summary, generalizing from multinomial logit to nested logit relaxes the I.I.A assumption, but not completely. This is an important but often overlooked point. Generalizing equation (29) by allowing both a_m

and s_m to vary buys no more in terms of the I.I.A. assumption; where I.I.A. was imposed across pairs with $a_m = a$ and $s_m = s \neq 1 \forall$, I.I.A. is still imposed.

For Example: Consider **The Model of Site Choice with Saltwater Sites, Lakes and Rivers,** where M has three elements; 1= saltwater fishing, 2=lake fishing; and 3=river fishing; $J_1=3$, $J_2=4$ and $J_3=5$. Where

$$F(<\epsilon_{mj}>) = \exp\{-(a_1[\exp(-s_1\epsilon_{11})+...+\exp(-s_1\epsilon_{13})]^{(1/s1)} + a_2[\exp(-s_2\epsilon_{21})+... +\exp(-s_2\epsilon_{24})]^{(1/s2)} + a_3[\exp(-s_3\epsilon_{31})+...+\exp(-s_3\epsilon_{35})]^{(1/s3)})\}$$

The ratio of the probabilities for any two sites of the same type will not be influenced by a change in the attributes of any other site or sites. For example, the ratio of probabilities for two of the rivers will not change if the attributes of any of the saltwater sites, lakes or other rivers change.

The ratio of the probabilities for any two sites that are different water types will not be influenced by a change in the attributes of site or sites that is not of one of these water types. For example, the ratio of probabilities for a lake site and a river site will not change if the attributes of one or more saltwater sites change.

The ratio of the probabilities for any two sites that are different water types will be influenced by a change in the attributes of a site or sites that is one of these water types. For example, the ratio of probabilities for a lake site and a river site will change if the attributes of some other lake or river site change.

II. Estimation

While this note is not about estimation per sec, a few comments about estimation are in order. The log of the likelihood function (examples are equation (21), (23) and (24)) can be maximized in one step by using a numerical algorithm to find the vector of parameter, $\{a_1, a_2,...,a_M; s_1, s_2,...,s_M$; and the parameters in the $<V_{mj}>$ functions$\}$ that maximize it. This approach is deemed

Full Information Maximum Likelihood, *FIML*.[11] Alternatively, one can adopt a two-stage sequential estimation, *SE*. The parameters in the V_{mj} for group m can be divided into two categories for purposes of estimation, those that just influence the allocation between alternatives in group m (that is, those that appear in the conditional probabilities, equation (19), for group m); and all other parameters. In the first step of sequential estimation, for each group one separately estimates just those parameters that determine the allocation amongst the alternatives in that group. This is done by maximizing the log of the likelihood function for the choice of each j in the group conditional on the choice of n. In this first stage, the model for each group is a one-level multinomial logit model with J_n alternatives. In the second stage of a sequential estimation, one estimates the $<a_m>$, $<s_m>$ and other parameters, given the parameter estimates from the first stage. While this two step estimation procedure is tempting, and is the easier approach given existing computer hardware and software, I recommend against it.

It has been known for a long time that the sequential technique will leads to parameter estimates that are not asymptotically efficient, and standard-error estimates that are inconsistent (Amemiya [1]). What was not known until recently was the degree of bias associated with parameter estimates obtained using the two-step sequential technique. Recent studies by Cameron [9] and [10], Hensher [19], Brownstone and Small [7], and Kling and Thompson [21] have found that this bias can be significant. Often sequential estimation generates parameter and welfare estimates that are substantially different from those generated by FIML (Kling and Thompson [21]). Software (e.g. Gauss and other such programs) to directly maximize the log of the likelihood function for the full model, equation (21), (23) or (24), are now widely available for both PCs and mainframes. There are numerous examples of FIML nested-logit estimation in both other fields and in

[11] Note that identification requires that one of the a_m is set to some positive constant; e.g. 1. If one's intent is to estimate a basic two-level nested model ($a_m = a$ and $s_m = s \; \forall$), the a cancels out and the parameter vector is just {s; and the parameters in the $<V_{mj}>$}.

recreational demand. See, for example Morey et. al. [28] and Morey et. al. [30], and the articles noted above.

Sometimes estimates from the sequential technique are advocated as starting value for the FIML estimation, but Hensher [19] even discourages that practice. I advocate the following. Obtain initial estimates by assuming that $a_m = 1$ and $s_m = 1 \forall M$; i.e. estimate a basic multinomial logit model assuming that there is no nest; that is initially assume all C alternatives are statistically independent. This will provide initial estimates for the parameters in each of the V_{mj} functions. Use these initial estimates as starting values to maximize the log of the likelihood function subject to the restrictions that $a_m = 1$ and $s_m = s \forall M$, with $s = 1$ for the starting value for s. If maximization is successful, the resulting parameter estimates are nested-logit FIML.

One caution is in order, estimation often leads to an estimate(s) for s (s_m) that is greater than zero but less than one. A result that violates McFadden's sufficient condition that s must be ≥ 1 for the nested logit model to be globally consistent with utility maximization.[12] McFadden [22] shows that **sufficient** conditions for the two-level nested logit model to be consistent with utility maximizing behavior are that $a_m > 0 \forall M$ and that $s_m \geq 1 \forall M$. These restriction guarantee that equation (8) is always positive for all $<\epsilon_{mj}>$ vectors ; a necessary condition for a function to be a density function. While a_m must be > 0 and s_m must be $> 0 \forall M$, Börsch-Supan [8] has shown that a nested-logit model where $0 < s_m < 1$ **can** be consistent with utility maximization if one weakens the restriction that equation (8) must always be nonnegative and only require that it be nonnegative for $<\epsilon_{mj}>$ vectors generated by the data set and parameter estimates. One must then check the more complicated necessary conditions identified by Börsch-Supan [8]. For more specifics on the specification (nesting structure) and estimation of NL models see Kling and Thompson [21].

[12]See footnote 8. One could impose the restriction that $s_m \geq 1$ by replacing, in estimation, s_m with $s_m = 1+e^{\theta_m}$

IV. Derivation of Expected Maximum Utility for the Two-level Nested Logit Model

The intent of this section is to use the basics of probability theory to derive expected maximum utility from equations (1) and (8). Section VI. uses expected maximum utility to derive a compensating variation and an equivalent variation. Before proceeding with the derivation of expected maximum utility from the nested -logit model it is important to point out that the expected maximum utility derived in this section is expected maximum utility *per choice occasion*, not expected maximum utility for the year or fishing season. One must remain cognizant of this if one's intent is to derive the per year compensating variation associated with a change in the attributes of a site or sites.

Let $\max(<U_{mj}>) \equiv \max(<V_{mj} + \epsilon_{mj}>)$ denote the largest element in the vector $<V_{mj} + \epsilon_{mj}>$. Therefore, given equation (1), expected maximum utility, U, is

$$(31)\quad U = \int_{\epsilon_{11}=-\infty}^{+\infty}\int_{\epsilon_{12}=-\infty}^{+\infty}\cdots\int_{\epsilon_{mj}=-\infty}^{+\infty}\cdots\int_{\epsilon_{MJ}=-\infty}^{+\infty}\max(V_{11}+\epsilon_{11},V_{12}+\epsilon_{12},\ldots,V_{mj}+\epsilon_{mj},\ldots,V_{MJ}+\epsilon_{MJ})$$

$$f(\epsilon_{11},\epsilon_{12},\ldots,\epsilon_{mj},\ldots,\epsilon_{MJ})d\epsilon_{JM}\ldots d\epsilon_{mj}\ldots d\epsilon_{12}d\epsilon_{11}$$

Equation (31) is the equation for the expected value of the function $\max(<V_{mj} + \epsilon_{mj}>)$. Equation (31) can be written more simply by dividing the density into C regions such that in region *ni* alternative *ni* is chosen (i.e., in region *ni*, alternative *ni* has maximum utility). Dividing into these *ni* regions, one obtains

$$(32)\quad U = \sum_{n=1}^{M}\sum_{i=1}^{J_n}\int_{\epsilon_{ni}=-\infty}^{+\infty}(V_{ni} + \epsilon_{ni})\, F_{ni}(<V_{ni}+\epsilon_{ni}-V_{mj}>)d\epsilon_{ni}$$

where, as noted in equation (7) $\int_{\epsilon_{ni}=-\infty}^{+\infty}F_{ni}(<V_{ni}+\epsilon_{ni}-V_{mj}>)d\epsilon_{ni} = Prob(ni)$. Equation (32) identifies maximum utility for **any** discrete choice model that is consistent with equation (1). In this sense equation (32) is quite general, and one could, in theory, plug any specific CDF, $F(<\epsilon_{mj}>)$ into equation (32) to derive the expected maximum utility associated with that CDF. A critical issue, as with the derivation of the Prob(ni), is when equation (32) will have a closed-form solution. It has a closed-form solution if one assumes the Generalized Extreme Value distribution denoted in equation (8), as is now demonstrated.

To obtain expected utility for the two-level nested logit model, substitute equation (10) into equation (32) to obtain

$$(33)\quad U = \sum_{n=1}^{M}\sum_{i=1}^{J_n}\int_{\epsilon_{ni}=-\infty}^{+\infty}(V_{ni}+\epsilon_{ni})\exp\{-\sum_{m=1}^{M} a_m[\sum_{j=1}^{J_m} e^{-s_m(V_{ni}+\epsilon_{ni}-V_{mj})}]^{(1/s_m)}\}$$

$$a_n[\sum_{j=1}^{J_n} e^{-s_n(V_{ni}-V_{nj}+\epsilon_{ni})}]^{(1/s_n)-1} e^{-s_n\epsilon_{ni}}\, d\epsilon_{ni}$$

Simplify, by making the change of variables $w = V_{ni} + \epsilon_{ni}$ ($\Rightarrow d\epsilon = dw$ because V_{ni} is a constant, and $\epsilon_{ni} = w - V_{ni}$), to obtain

$$(34)\quad U = \sum_{n=1}^{M}\sum_{i=1}^{J_n}\int_{w=-\infty}^{+\infty} w\,\exp\{-\sum_{m=1}^{M} a_m[\sum_{j=1}^{J_m} e^{-s_m(w\,-\,V_{mj})}]^{(1/s_m)}\}$$

$$a_n[\sum_{j=1}^{J_n} e^{-s_n(w\,-\,V_{nj})}]^{(1/s_n)-1} e^{-s_n(w\,-\,V_{ni})}\, dw$$

Note that the term in equation (34), $\exp\{-\sum_{m=1}^{M} a_m[\sum_{j=1}^{J_m} e^{-s_m(w\,-\,V_{mj})}]^{(1/s_m)}\}$

$= \exp[-De^{-w}]$

where

$$(35)\quad D = \sum_{m=1}^{M} a_m [\sum_{j=1}^{J_m} e^{s_m V_{mj}}]^{(1/s_m)}$$

and that the term in equation (34), $[\sum_{j=1}^{J_n} e^{-s_n(w - V_{nj})}]^{(1/s_n)-1}$

$$= e^{-w} e^{s_n w} [\sum_{j=1}^{J_n} e^{s_n V_{nj}}]^{(1/s_n)-1}.$$

> **For example: In the model of participation and site choice**, the season is divided into a finite number of periods (choice occasions) so U is expected maximum utility per choice occasion. Expected maximum utility for the season is the sum of the per-period U's. If attributes vary from period to period, U will vary across the periods.
>
> In the **model of site choice with saltwater sites, lakes and rivers**, each trip is a choice occasion so U is expected maximum utility **per-trip**. Per-trip expected maximum utility does not easily translate into seasonal expected maximum utility.
>
> As we will see when it comes to the derivation of compensating variation, the distinction between per-period and per-trip expected maximum utility is quite important.

Substituting these two simplifications into equation (34) and moving all of the terms that do not contain w, with the exception of D, to the left of the integral sign, one obtains

$$(36)\quad U = \{\sum_{n=1}^{M} \sum_{i=1}^{J_n} a_n e^{s_n V_{ni}} [\sum_{j=1}^{J_n} e^{s_n V_{nj}}]^{(1/s_n)-1}\}$$

$$\int_{w=-\infty}^{+\infty} w e^{-w} \exp\{-De^{-w}\} dw$$

Now examine the first term in equation (36). It equals D. This follows from Euler's Theorem, as is now demonstrated. D, equation (35), is homogenous of degree one in $(e^{V_{ni}})$, and

$$(37)\quad \frac{\partial D}{\partial (e^{V_{ni}})} = e^{-V_{ni}} a_n e^{s_n V_{ni}} [\sum_{j=1}^{J_n} e^{s_n V_{nj}}]^{(1/s_n)-1}$$

Therefore, by Euler's Theorem[13],

$$(38)\quad \sum_{n=1}^{M}\sum_{i=1}^{J_n} e^{V_{ni}}\frac{\partial D}{\partial(e^{V_{ni}})} = \sum_{n=1}^{M}\sum_{i=1}^{J_n} e^{V_{ni}}\{e^{-V_{ni}}a_n e^{s_n V_{ni}}[\sum_{j=1}^{J_n} e^{s_n V_{nj}}]^{(1/s_n)-1}\}$$

$$= \sum_{n=1}^{M}\sum_{i=1}^{J_n} a_n e^{s_n V_{ni}}[\sum_{j=1}^{J_n} e^{s_n V_{nj}}]^{(1/s_n)-1} = D$$

Since the first term in Equation (36) equals D, equation (36) implies

$$(39)\quad U = \int_{w=-\infty}^{+\infty} wDe^{-w}\exp\{-De^{-w}\}dw$$

This is where things get exciting. Note that the density function for a univariate Extreme Value Distribution with scale parameter D > 0 is

$$(40)\quad f(w) = De^{-w}\exp\{-De^{-w}\} \qquad \text{(Johnson And Kotz [20])}$$

Therefore, since $E(w) = \int_{w=-\infty}^{+\infty} wf(w)dw$,equation (39) is the expected value (mean) of an Extreme Value Distribution with scale parameter D. It is well known that this expected value is lnD + .5772..., where .57721... is Euler's Constant. Therefore expected maximum utility is

$$(41)\quad U = \ln D + .57721....$$

[13]Euler's Theorem states $D = \sum_{n=1}^{M}\sum_{i=1}^{J_n} \alpha_{ni}\frac{\partial D}{\partial \alpha_{ni}}$ if D is homogenous of degree one in the α_{ni} . In our case, $\alpha_{ni} = e^{V_{ni}}$

Equation (41) is used in Section VI. to derive a compensating variation and an equivalent variation from the nested logit model. Remember that U is expected maximum utility per choice occasion.

As an aside, note that it can shown that

$$(42) \quad \frac{\partial U}{\partial V_{ni}} = Prob(ni)$$

Said loosely, equation (42) is the discrete-choice random-utility analog of Roy's Identity where V_{ni} is interpreted as the negative of the *normalized price* of alternative ni.

V. The Functional Form of the Conditional Indirect Utility Functions, $<V_{mj}>$, and Budget Exhaustion

Let I denote the budget, p_{mj} the price (cost) of alternative *mj*, and β^{mj} the vector of other observed attributes of alternative *mj*. View the choice set as containing C mutually exclusive alternatives and a numeraire composite good. During each choice occasion, the individual is constrained to consume one, but only one, of the C alternatives, and then spends the rest of his or her budget on the numeraire good. For example, if alternative *ni* is chosen, $(I - p_{ni})$ is spent of the numeraire good and budget exhaustion requires that V_{mj} is restricted to functions of the form

$$(43) \quad V_{mj} = V_{mj}((I - p_{mj}), \beta^{mj})$$

It is common to assume that equation (43) is linear in its parameters, but this is both restrictive and unnecessary. If one restricts equation (43) to be linear in $(I - p_{ni})$ of the form

$$(43a) \quad V_{mj} = \mu(I - p_{mj}) + h_{mj}(\beta^{mj})$$

, a commonly imposed restriction, μ is the constant marginal utility of money, and the budget, I, is **not** a determinant of which alternative is chosen (that is, when equation (43a) is adopted, the budget does not appear as a variable in the probability equations). Nested logit models that assume equation (43a) are deemed NL models *with zero income effects*. As we'll soon see, one only gets closed-form solutions for the compensating, and equivalent, variation, if one assumes (43a).

However, this lack of closed-form solutions is not a big deal because, given equation (43) and its parameter values, it is easy to derive an exact compensating variation and an exact equivalent variation for any specific change in prices and/or other attributes.

The reason that the budget drops out of the choice probabilities when constant marginal utility of money is assumed is that any attribute (budget, price or other attribute) that has a constant magnitude across all the alternatives, and whose coefficients in the conditional indirect utility functions do not vary across the alternatives, will not influence the choice of alternative (μI is just one example).

More generally, any attribute that has a constant magnitude across some subset of the alternatives, and whose coefficients do not vary across the alternatives in this subset, will influence whether this subset is chosen but not influence which alternative in this subset is chosen. For example, if

For Example, In the **model of site choice with saltwater sites, lakes and rivers**, expected catch rate is likely an important attribute of all of the sites. Assume, for example, that all lake sites have the same expected catch rate and the coefficient on catch is the same constant in all of the conditional indirect utility functions for lake sites, but not the same constant for other sites. In this case, expected catch rate will influence whether one chooses a lake site, but, conditional on a lake trip being chosen, will not influence which specific lake is visited.

(44) $V_{mj} = \mu_m(I - p_{mj}) + h_{mj}(\beta^{mj})$

the budget, I, will influence whether an alternative of type m will be chosen, but not influence which alternative of type m will be chosen. If size is an attribute, all alternatives of type *m* are size 10, and the coefficient on size is the same in all of the conditional indirects for alternatives of type m, conditional on you choosing an alternative of type m, size will not influence which of the alternatives of type *m* you will choose. This can be confirmed by examining equation (19).

VI. A Compensating Variation and An Equivalent Variation

Let $P \equiv <p_{mj}>$, $\beta \equiv <\beta^{mj}>$, and consider a change from the initial state $\{I^o, P^o, \beta^o\}$ to some proposed state $\{I^1, P^1, \beta^1\}$. For each choice occasion , define the compensating variation, CV, associated with this change as the monetary compensation (or payment) in the proposed state that would make the expected maximum utility in the proposed state equal to the maximum expected utility in the initial state; that is,

(45)[14] $U^0 \equiv U(I^0, P^0, \beta^0) = U(I^1 - CV, P^1, \beta^1)$ where

expected maximum utility, U, is defined by equations (41) and (35). Remember that CV is the compensating variation per choice occasion, not for the year or season. At the end of this section, I will discuss whether and, if possible, how the seasonal compensating variation can be derived from the per choice-occasion compensating variation.

Define an equivalent variation per choice occasion , EV, as

(46) $U^1 \equiv U(I^1, P^1, \beta^1) = U(I^0 + EV, P^0, \beta^0)$

i.e., the monetary compensation (or payment) in the initial state that would make expected maximum utility in the initial state equal to expected maximum utility in the proposed state.

If one imposes zero income effects (equation (43a)), equation (45) can be solved for CV, and CV = EV. Specifically,

---

[14]Note that this definition for the compensating variation is the random-utility analog of the standard definition of compensating variation in terms of a deterministic utility function. The CV generated by equation (45) is the compensation (or payment) in the proposed state that would make an **observer's** expectation of maximum utility in the proposed state equal to his expectation of expected maximum utility in the initial state. Hanemann [18] notes two other possible definitions of the compensating variation. We have defined U as expected maximum utility from the observer's perspective. Define u as actual utility. The compensating variation from the individual's perspective, $\hat{cv}$, is not a random variable. It is $u(y^o, P^o, \beta^o, <\epsilon_{mj}{}^o>) = u(y^1 - \hat{cv}, P^1, \beta^1, <\epsilon_{mj}{}^1>)$. $\hat{cv}$ is a random variable from the observer's perspective, and one might define a compensating variation, $\hat{CV}$, as the observer's expectation of $\hat{cv}$. Hanemann also defines a compensating variation, $\bar{CV}$, as the compensation (or payment) in the proposed state that would make the probability that utility in the proposed state (with the compensation) is greater than utility in the initial state 50%. Whereas $\hat{CV}$ is the mean of the distribution of the $\hat{cv}$, $\bar{CV}$ is the median. As Hanemann [18] demonstrates, for the nested logit model, CV = $\bar{CV}$, but CV does not, in general, equal $\hat{CV}$. CV = $\hat{CV}$ in the nested logit model only if one imposes zero income effects.

(47) $CV = EV = (1/\mu)[U(I^0, P^1, \beta^1) - U(I^0, P^0, \beta^0)] + (I^1 - I^0)$

$$= (1/\mu)[\ln D^1 - \ln D^0]$$

where D is defined in equation (35), D^1 is D evaluated at I^0, P^1 and β^1, and D^0 is D evaluated at I^0, P^0 and β^0. Intuitively, in the case of zero income affects, the CV is just the change in expected maximum utility converted into a money metric by multiplying by the inverse of the constant marginal utility of money, $(1/\mu)$.

When one does not impose zero income effects, CV $\neq$ EV, equation (45) cannot be solved to obtain a closed form solution for the CV and equation (46) cannot be solved to obtain a closed form solution for the EV. However, it is still easy to numerically calculate an individual's CV, or EV, for any proposed change. For example, given the estimated parameter values, an individual's CV for any proposed change can be calculated by using any numerical minimization algorithm to find the CV that minimizes $[U(y^0, P^0, \beta^0) - U(y^1 - CV, P^1, \beta^1)]^2$

where U is defined by equation (41).[15]. The CV that minimizes this expression will be the individual's CV for the change.

Remember that CV is not the per year (or per season) compensating variation associated with a change in the attributes of the alternatives, but rather the compensating per choice occasion. If the year is divided into a fixed number of choice occasion (e.g. weeks) and during each of the periods the nested logit model includes all possible alternatives in the choice set, the compensating variation for the year (season) is easily obtained by multiplying the per-period CV by the number of periods in the year (season).[16]

[15]For example one can use the *optimization* routine in the statistical package *Gaussi* or the *FindMinimum* command in the mathematical software *Mathematica*.

[16]Note that alternatives can be grouped. For example, the complete set of alternatives facing the individual could be lumped in two categories, fishing trips and nonfishing trips where nonfishing trips include staying at home, going bowling, etc.

Alternatively, if the choice set for each choice occasion is restricted, things are more complicated and one cannot get the compensating variation for the year by simply multiplying the compensating variation per choice occasion by the number of choice occasions in the year. The problem is that when the choice set is restrictived, the number of choice occasions in the year becomes a function of the attributes of the alternatives in the restrictive choice set. Multiplying the CV by the number of choice occasions in the initial state, $\{I^o, P^o, \beta^o\}$, provides only a lower bound on the compensating variation for the year. Multiplying the CV by the number of choice occasions in proposed state, $\{I^1, P^1, \beta^1\}$, provides only an upper bound on the compensating variation for the year. Neither is necessarily a close approximation. Details are provided in Morey [27].

> **For example:** In the **model of participation and site choice**, the year is divided into a finite number of periods and the individual is not constrained to fish. In this case, CV is the per-period compensating variation and the compensating variation for the year is easily obtained by multiplying the CV by the number of periods.
>
> In the **model of site choice with saltwater sites, lakes and rivers**, choice occasions are trips and on each trip the individual is constrained to fish (staying home is not an option). In this case, CV is the compensating variation per trip. Define trips^0 as the predicted number of trips in the initial state and trips^1 as the predicted number of trips in the proposed state. CV multiplied by trips^0 approximates the compensating variation from below. CV multiplied by trips^1 approximates the compensating from above.

VII. Expanding The Nest

The nested logit model can be expanded into as many levels as one desires. For a three-level nest, equation (1) expands to

(1-3level) $U_{lmj} = V_{lmj} + \epsilon_{lmj} \qquad \forall\ (lmj) \in C$

where the l dimension of the choice set is characterized in term of L distinction types. As before, the m dimension has M distinct types and the j dimension J distinct types. To generate a simple three-level nested-logit model where $s_m = s \; \forall \; m$ and $a_m = 1 \forall \; m$ assume[17]

$$\text{(8-3level-1)} \quad F(<\epsilon_{lmj}>) = \exp\{-\sum_{l=1}^{L} [\sum_{m=1}^{M} [\sum_{j=1}^{J_m} e^{-s\epsilon_{lmj}}]^{(t/s)}]^{1/t}\}$$

The derivation of the Prob(gni) and expected maximum utility follow the same logic as in the two-level case, the notation is just messier. If one bashes through the derivations, one obtains

$$\text{(16-3level-1)} \quad Prob(gni) = \frac{e^{sV_{gni}} [\sum_{m=1}^{M_g} [\sum_{j=1}^{J_m} e^{sV_{gmj}}]^{(t/s)}]^{(1/t)-1} [\sum_{j=1}^{J_n} e^{sV_{gnj}}]^{((t/s)-1)}}{\sum_{l=1}^{L} [\sum_{m=1}^{M} [\sum_{j=1}^{J_m} e^{sV_{lmj}}]^{(t/s)}]^{1/t}}$$

and per-choice occassion expected maximum utility is

$$\text{(41)} \quad U = \ln D + .57721....$$

where

$$\text{(35-3level-1)} \quad D = \sum_{l=1}^{L} [\sum_{m=1}^{M} [\sum_{j=1}^{J_m} e^{sV_{lmj}}]^{(t/s)}]^{1/t}$$

[17]A sufficient, but not necessary, condition for this density function to be well behaved is $s \geq t \geq 1$.

One Example of a three-level nest: A model of participation and site choice for Atlantic Salmon fishing (Morey et al [28]) with C = 9 alternatives (staying at home and visiting one of eight salmon rivers). Assume L has two elements; 1= going salmon fishing and 0 = not salmon fishing. M_1 has two element; 1= fishing a river in Maine and 2 = fishing a river in Canada. J_1 includes five rivers in Maine and J_2 includes three rivers in Canada. A three-level nest seems appropriate. One would expect the random components in the conditional indirect utility functions for the Maine rivers to be more correlated with each other than they are with the random components in the conditional indirect utility functions for the Canadian rivers. And, one would expect the random components in the conditional indirect utility functions for the rivers to be more correlated with each other than they are with the random component in the conditional indirect utility functions for nonparticipation.

In which case, equation (8-3level) takes the form

$$F(<\epsilon_{lmj}>) = \exp\{-\exp(-\epsilon_0) - [\,[\exp(-s\epsilon_{11})+...+\exp(-s\epsilon_{15})]^{t/s} + [\exp(-s\epsilon_{21})+... \exp(-s\epsilon_{23})]^{t/s}\,]^{1/t}\}$$

and the probability of choosing the first river in Canada is

$$\text{Prob}(121) = \exp(sV_{121})\,[\,[\exp(sV_{111})+...+\exp(sV_{115})]^{t/s} + [\exp(sV_{121})+...+\exp(sV_{123})]^{t/s}\,]^{1/t-1} \text{ multiplied by } [\exp(sV_{121})+...+\exp(sV_{123})]^{t/s-1}$$

and divided by

$$\{\exp(V_0) + [\,[\exp(sV_{111})+...+\exp(sV_{115})]^{t/s} + [\exp(sV_{121})+...+\exp(sV_{123})]^{t/s}\,]^{1/t}$$

```
@ nstl-3lv.cmd  June 30, 92 @

@ this is the gaussi program that was used to estimate the three level NL model of Atlantic
Salmon fishing that was described in the previous box. For more details see Morey et al [28].
@

library maxlik; #include maxlik.ext; maxset;

   dataset = "mst4lgt";        output file = nstl-3lv.out on;

proc li(b,x);     @ "li" is the procedure that generates the ln of the lik function. "li" is called
below by the maximum likelihood application, "maxlik" @

local ppy, evp, evd, evm, evk, evs, evns, evnb, evq, evn, vp, vm, vd, vk, vs, vns, vnb, vq, vn,
        inclusm, inclusc, inclusp, inclus, linclus, lsump, lmsum, lcsum, x;

ppy = x[.,4]/50; @ x[.,4] is income @

@ the following are the exp of the conditional indirects with the conditional indirects for the
site alternatives multiplied by s;i.e b[11]. Note that the conditional indirect for
nonparticipation, n, is not multiplied by s because it cancels out. @

@ In the following b[4] is a constant term, b[11]=s, b[12] = t, x[.,1] = years salmon fishing.
x[.,2] = 1 if member of a fishing club and zero otherwise, x[.,3] = individual's age, x[.,5] =
number of days individual did not salmon fish,  x[.,6]- x[.,13] are the expected catch rates at
the eight sites, x[.,14] -x[.,21]are trip costs for the eight sites,and x[.,22]- x[.,29] are the number
of trips each individual took to each of the eight sites, @

 evp = exp(b[11]*(b[1]*(ppy-x[.,14])+b[2]*x[.,6]+b[3]*x[.,6]^.5
    +b[10]*((1728.4720+ppy-x[.,14])^.5)));

evm = exp(b[11]*(b[1]*(ppy-x[.,15])+b[2]*x[.,7]+b[3]*x[.,7]^.5
    +b[10]*((1728.4720+ppy-x[.,15])^.5)));

evd = exp(b[11]*(b[1]*(ppy-x[.,16])+b[2]*x[.,8]+b[3]*x[.,8]^.5
    +b[10]*((1728.4720+ppy-x[.,16])^.5)));

evk = exp(b[11]*(b[1]*(ppy-x[.,17])+b[2]*x[.,9]+b[3]*x[.,9]^.5
    +b[10]*((1728.4720+ppy-x[.,17])^.5)));

evs = exp(b[11]*(b[1]*(ppy-x[.,18])+b[2]*x[.,10]+b[3]*x[.,10]^.5
    +b[10]*((1728.4720+ppy-x[.,18])^.5)));

 @ the program is continued in the next box @
```

```
@ nstl-3lv.cmd   June 30, 92  continuation of program from previous example box  @

evnb= exp(b[11]*(b[1]*(ppy-x[.,20])+b[2]*x[.,12]+b[3]*x[.,12]^.5
     +b[10]*((1728.4720+ppy-x[.,20])^.5)));

evq = exp(b[11]*(b[1]*(ppy-x[.,21])+b[2]*x[.,13]+b[3]*x[.,13]^.5
     +b[10]*((1728.4720+ppy-x[.,21])^.5)));

@the next line is the exp of the condit indirect for nonpartic @

evn = exp(b[1]*ppy+b[4]+b[5]*x[.,1]+b[6]*x[.,2]+b[7]*x[.,3]
     +b[8]*x[.,1]^.5+b[9]*x[.,3]^.5+b[10]*((1728.4720+ppy)^.5));

vp=ln(evp); vm=ln(evm); vd=ln(evd); vk=ln(evk); vs=ln(evs);vns=ln(evns); vnb=ln(evnb);
vq=ln(evq); vn=ln(evn);

@ Note b[12] is t @

inclusm = (evp+evm+evd+evk+evs)^(b[12]/b[11]);
inclusc = (evns+evnb+evq)^(b[12]/b[11]);
inclusp = (inclusm + inclusc)^(1/b[12]);

@ inclus is the denomin in all the prob @

inclus  = evn + inclusp;
linclus=ln(inclus);
lsump = ((1/b[12])-1)*ln(inclusm+inclusc);
lmsum = ((b[12]/b[11])-1) * ln(evp+evm+evd+evk+evs);
lcsum = ((b[12]/b[11])-1) * ln(evns+evnb+evq);

@ the next command calculates the contribution to the log of the lik function for each indiv in
the sample @

retp(x[.,22].*(vp + lsump + lmsum - linclus)
     + x[.,23].*(vm + lsump + lmsum - linclus)
     + x[.,24].*(vd + lsump + lmsum - linclus)
     + x[.,25].*(vk + lsump + lmsum - linclus)
     + x[.,26].*(vs + lsump + lmsum - linclus)
     + x[.,27].*(vns+ lsump + lcsum - linclus)
     + x[.,28].*(vnb+ lsump + lcsum - linclus)
     + x[.,29].*(vq + lsump + lcsum - linclus)
     + x[.,5].*(vn - linclus));

endp;     @ this is the end of the procedure "li" @

 @ the program is continued in the next box @
```

```
@ nstl-3lv.cmd   June 30, 92  continuation of program from previous example box  @

@ the following are the converg values from the 6/30/92 run - the AJAE estimates @

startv = {.002190, -1.729160, 5.912200, 8.850202, .095329, -.805250,.170146, -1.227537,
-1.856089, 1.088725, 1.307125, .611724};

_title = "NSTL-3LEV INCOME EFFECTS ";
_mlmiter = 2000;
_mlgtol  = .0001;
{bbb,f0,g,h,retcode}=maxprt(maxlik("mst4lgt",0,&li,startv));

bmm94 = bbb;     save bmm94;

 output off;
```

Consider now a more general three-level nested-logit model where[18]

$$\text{(8-3level-2)} \quad F(<\epsilon_{lmj}>) = \exp\{-\sum_{l=1}^{L} x_l \, [\sum_{m=1}^{M} a_m \, [\sum_{j=1}^{J_m} e^{-s_m \epsilon_{lmj}}]^{(t_l/s_m)}]^{1/t_l}\}$$

$$\text{(16-3level-2)} \quad Prob(gni) = \frac{e^{s_n V_{gni}} \, [\sum_{m=1}^{M_g} [\sum_{j=1}^{J_m} e^{s_m V_{gmj}}]^{(t_g/s_m)}]^{(1/t_g)-1} \, [\sum_{j=1}^{J_n} e^{s_n V_{gnj}}]^{((t_g/s_n)-1)}}{\sum_{l=1}^{L} [\sum_{m=1}^{M} [\sum_{j=1}^{J_m} e^{s_m V_{lmj}}]^{(t_l/s_m)}]^{1/t_l}}$$

(16-3level-2a)

$$= \frac{e^{s_n(\chi_g + \alpha_n + V_{gni})} \, [\sum_{m=1}^{M_g} [\sum_{j=1}^{J_m} e^{s_m(\chi_g + \alpha_m + V_{gmj})}]^{(t_g/s_m)}]^{(1/t_g)-1} \, [\sum_{j=1}^{J_n} e^{s_n(\chi_g + \alpha_n + V_{gnj})}]^{((t_g/s_n)-1)}}{\sum_{l=1}^{L} [\sum_{m=1}^{M} [\sum_{j=1}^{J_m} e^{s_m(\chi_l + \alpha_m + V_{lmj})}]^{(t_l/s_m)}]^{1/t_l}}$$

where $\alpha_m = \ln(a_m)$, and $\chi_l = \ln(x_l)$. The parameter χ_l adds a group specific constant term to each of the V_{lmj}. As noted on page 9, the α_m add a group specific constant term to each of the V_{lmj}

$$\text{(35-3level-2)} \quad D = \sum_{l=1}^{L} x_l \, [\sum_{m=1}^{M} a_m \, [\sum_{j=1}^{J_m} e^{s_m V_{lmj}}]^{(t_l/s_m)}]^{1/t_m}$$

[18]Sufficient, but not necessary, conditions for this density function to be well behaved are $a_m > 0 \; \forall \; m$, $x_l > 0 \; \forall \; l$; and $\{s_m \geq t_l \geq 1 \;\; \forall \; m \in M_l\}$, $l=1,2,...L$.

References

1. T. Amemiya, "On a Two-Step Estimation of a Multivariate Logit Model," *Journal of Econometrics*, **8**, 12-21 (1978).

2. M. Ben-Akiva, and S.R. Lerman, "Discrete Choice Analysis: Theory and Application to Travel Demand," MIT Press, Cambridge (1985).

3. N.E. Bockstael, W.M. Hanemann and C.L. Kling, "Modelling Recreational Demand in a Multiple Site Framework," *Water Resources Research* **23**, 951-60 (1987).

4. N.E. Bockstael, W.M. Hanemann and I.E. Strand, "Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models," Report prepared for the Environmental Protection Agency under Cooperative Agreement CR-811043-01-0, Washington D.C. (1986).

5. N.E. Bockstael, K.E. McConnell and I.E. Strand, "A Random Utility Model of Sport Fishing: Some Preliminary Results for Florida," *Marine Resource Economics*, **6**, 245-260 (1989).

6. N.E. Bockstael, K.E. McConnell and I.E. Strand, Recreation, in "Measuring the Demand for Environmental Quality," (John B. Braden and Charles Kolstad, eds.), New York, North Holland (1991).

7. D. Brownstone and K.A. Small, "Efficient Estimation of Nested Logit Models," *Journal of Business and Economic Statistics*, **7**, 67-74 (1989).

8. A. Börsch-Supan, "On the Compatibility of Nested Logit Models with Utility Maximization," *Journal of Econometrics*, **43**, 373-388 (1990).

9. T.A. Cameron, Qualitative Choice Modeling of Energy Conservation Decisions, unpublished Ph.D. Dissertation, Princeton University, Department of Economics (1982).

10. T.A. Cameron, "A Nested Logit Model of Energy Conservation Activity by Owners of Existing Singe Family Dwellings," *The Review of Economics and Statistics*, **55**, 205-211 (1985).

11. R. Carson, W.M. Hanemann and T Wegge, "Southcentral Alaska Sport Fishing Study," Report prepared by Jones and Stokes Associates for Alaska Department of Fish and Game, Anchorage, AK 99502 (1987)

12. P. Caulkins, R. Bishop and N. Bouwes, "The Travel Cost Model for Lake Recreation: A Comparison of Two Methods for Incorporating Site Quality and Substitution Effects," *American Journal of Agricultural Economics*, **68**, 291-297 (1986).

13. M. Creel and J. Loomis, "Recreation Value of Water to Wetlands in the San Joaquin Valley: Linked Multinomial Logit and Count Data Trip Frequency Models," *Water Resources Research*, **28**, 2597-2606 (1992).

14. D. Feenburg and E. Mills, "Measuring the Benefits of Water Pollution Abatement," Academic Press, New York (1980).

15. C.A. Jones and Y.D. Sung, "Models for Censored Duration Data, Competing Risks, and Time-Varying Covariates: with an Application to Recreational Trip Demand," discussion paper, NOAA, US Department of Commerce (1992).

16. J.A. Hausman, G.K. Leonard and D McFadden, "A Utility-Consistent, Combined Discrete Choice and Count Data Model: Assessing Recreational Use Losses Due to Natural Resource Damage," discussion paper, Department of Economics, MIT, Cambridge, MA (1993).

17. W. M. Hanemann, "A Methodological and Empirical Study of the Recreation Benefits from Water Quality Improvements," PhD dissertation, Department of Economics, Harvard University (1978).

18. W. M. Hanemann, "Welfare Analysis with Discrete Choice Models," working paper, Department of Resource and Agricultural Economics, University of California, Berkeley, (March 1985).

19. D.A. Hensher, "Sequential and Full Information Maximum Likelihood Estimation of a Nested Logit Model," *The Review of Economics and Statistics*, **56**, 657-667 (1986).

20. N.L Johnson, and S. Kotz, "Distributions in Statistics - Continuous Univariate Distributions. Vols. 1 and 2," Wiley, New York (1970).

21. C.L. Kling and C.J. Thomson, "The Implications of Model Specification for Welfare Estimation in Nested Logit Models," paper presented at the W-133 meetings, Tucson, (February 1994).

22. D. McFadden, Modelling the Choice of Residential Location, chapter 3 in "Spatial Interaction Theory and Planning Models" (A. Karlqvist, L, Lundqvist, F. Snickars and J Weibull, Eds.) North-Holland, Amsterdam (1978).

23. D. McFadden, Econometric Models of Probabilistic Choice, chapter 5 in "Structural Analysis of Discrete Data with Econometric Applications" (C. Manski and D. McFadden eds.), MIT Press, Cambridge (1981).

24. J.W. Milon, "A Nested Demand Shares Model of Artificial Marine Habitat Choice by Sport Anglers," *Marine Resources Research*, **5**, 191-213 (1988).

25. E.R. Morey, "The Demand for Site-Specific Recreational Activities: A Characteristics Approach," *Journal of Environmental Economics and Management*, 8, 245-271 (1981).

26. E.R. Morey, "Derivation of Expected Consumer's Surplus from a Nested-Logit Model of Consumer Choice," Discussion Paper, Department of Economics, University of Colorado, Boulder, CO 80309-0256, February 1989.

27. E.R. Morey, "What is Consumer's Surplus Per Day of Use, When is it a Constant Independent of the Number of Days of Use, and What Does It Tell Us About Consumer's Surplus," *Journal of Environmental Economics and Management* **26**, forthcoming 1994.

28. E.R. Morey, R.D. Rowe and M. Watson, "A Repeated Nested-Logit Model of Atlantic Salmon Fishing," *American Journal of Agricultural Economics*, **75**, 578-592 (August 1993).

29. E.R. Morey, W.D. Shaw and R.D. Rowe, "A Discrete Choice Model of Recreational Participation, Site Choice, and Activity Valuation When Complete Trip Data Are Not Available," *Journal of Environmental Economics and Management*, **20**,181-201 (1991).

30. E.R. Morey and R.D. Rowe with the assistance of B. Breffle and D. Shaw, "Assessment of Damages to Anglers and Other Recreators From Injuries to the Upper Clark Fork River Basin," Report prepared for the State of Montana, Natural Resources Damage Program, December 1993.

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

32. G.R. Parsons and M.J. Kealy, "A Demand Theory for Number of Trips in a Random Utility Model of Recreation," working paper #94-1, Department of Economics, University of Delaware (1994).

33. G.R. Parsons and M.S. Needleman, "Site Aggregation in a Random Utility Model of Recreation," *Land Economics*, **68**, 418-433 (1992).

The Implications of Model Specification for Welfare Estimation in Nested Logit Models

by

Catherine L. Kling
Department of Economics
Iowa State University
Ames, Iowa 50011

and

Cynthia J. Thomson
National Marine Fisheries Service
Southwest Fisheries Center
La Jolla, California 92038

May 13, 1994

Acknowledgements: The authors have benefitted from comments of participants at the W-133 regional research project. The results presented here are preliminary and subject to revision.

The Implications of Model Specification for Welfare Estimation in Nested Logit Models

Discrete choice models of the demand for environmental goods have been used to value changes in environmental quality or environmental experiences beginning with work by Hanemann (1978), Feenburg and Mills (1980), and Caulkins, Bishop, and Bouwes (1986). Recently, discrete choice models have seen even greater use (Bockstael, Hanemann, and Kling (1987), Bockstael, McConnell, and Strand (1989); Creel and Loomis (1992); Jones and Stokes Associates (1987); Milon (1988); and Morey, Shaw, and Rowe (1991); Morey, Rowe, and Watson (1993)) as increased computational capacity has allowed the estimation of these models with large data sets and many choice alternatives.

In these models, individuals are assumed to choose to visit the site that yields the highest utility per choice occasion. Assuming that each individual has a additive error that follows an extreme value distribution yields a logit model. The benefits of environmental improvement or preservation are computed based on the estimated coefficients. The main strength of the logit discrete choice model is its ability to model an individual's choices among a range of quality differentiated alternatives. This is well suited to many cases of recreation demand where environmental quality is an important component of choice. An important disadvantage of the simple logit model is the now famous independence of irrelevant alternatives (IIA) problem which implies that the addition of a new alternative does not change the choice probabilities between any two existing alternatives. One solution to the IIA problem is to use a generalized extreme value distribution which allows grouping of the alternatives and, therefore, correlation among the errors within each group. These models are generally referred to as nested logit models. [1]

The use of the nested logit specification for welfare evaluation raises several important modelling issues that have received little systematic analysis. First, the researcher must choose a nesting structure: which groups of alternatives should be grouped together to allow correlation between alternatives within that group. For example, suppose a recreationist chooses from among three activities, recreational boating, fishing, or picnicking, at three different locations (sites). There are thus nine activity/site alternatives. A

[1] See Morey (1992) for a nice derivation of the nested logit model and an explanation of the relationship between the nested model and the generalized extreme value model.

multinomial logit model with no nesting structure would treat the individual as choosing from among nine alternatives with the IIA property holding between all nine choices. One possible nesting structure is to group each of the three sites together for each activity. IIA would still hold within the group, i.e., between each site for each activity, but not between sites over different activities. An alternative grouping would be to group the three activities together for each site. In this case, IIA would hold within the activities for each site, but not between sites. There is no theoretical basis upon which researchers can choose between various nesting structures.

One common interpretation of the nesting structures has been a sequential decision making process. For the first grouping structure described above, the interpretation would be that the individual first chooses which activity to undertake and then which site to visit to undertake it. The second grouping would reverse the order of the decision: first the site would be chosen and then the activity. Although appealing in interpreting some correlation patterns, the sequential interpretation is not necessary. In fact, all that the grouping patterns do is introduce correlation among choices. This correlation could be introduced by individuals undertaking decisions in a sequential way, but these correlations may be introduced for a variety of other reasons such as omitted variables. Alternatives are correlated from the perspective of the researcher when there are similar unobserved factors between alternatives. As Train, McFadden, and Ben-Akiva (1987) note, these correlations arise "from patterns in the researcher's lack of information, rather than from the households' decision processes (page 113)."

Alternative decision structures imply different tree structures and consequently, different coefficient and welfare estimates. Although it is likely that the structure of the decision process is critical to the resulting welfare measures, almost no sensitivity analysis has been performed on the implications of alternative structures in the literature on environmental benefits.

A second issue in the estimation of nested logit models is the choice of estimator. In the context of recreation demand, estimation of such models has typically been accomplished in a sequential fashion, estimating first the probability of choosing an alternative within a group and then the probability of choosing each group.

Recent work in applied econometrics suggests that the sequential estimation methods used in most of these prior studies is asymptotically inefficient and yields inconsistent standard error estimates in the second stage of estimation. Further, in many applications the coefficient estimates from sequential estimation appear quite different from simultaneous estimation and several authors recommend a simultaneous estimation strategy (Cameron (1985); Hensher (1986); Borsch-Supan (1987); Brownstone and Small (1989)). Morey (1994) concludes on the basis of their findings that sequential estimation should be avoided. The simultaneous strategy is termed full-information maximum likelihood (FIML). Previous work, primarily on transportation problems, has focussed on the effect of FIML vs sequential estimation on coefficient estimates. However, the implications of estimator choice for welfare estimation has not been previously examined.

Additional issues arising in model specification and estimation have to do with the choice of variables included in the model and whether coefficient values are constrained to be equal across levels of the nest. It is possible to estimate nested logit models without constraining the coefficients on the variables to be the same across the groups of alternatives. Perhaps more interestingly, it is also possible to estimate separate dissimilarity parameters for each group which measure the degree of substitution within and between groups. These generalizations have rarely been empirically implemented.

Though there has been limited work examining the importance of nesting order, choice of variables, or equivalence of coefficients across nesting levels, some important work on specification issues in discrete choice models when applied to environmental valuation have been undertaken. Parsons and Needelman (1992) examine the implications of site aggregation and Parsons and Kealy (1992) examine the effects of random opportunity sets on estimation results and welfare measurement. Morey, Rowe, and Watson (1992) investigate the effects of including income effects and adopting a nested logit structure rather than a non-nested model.

Given the policy significance of estimating environmental benefits accurately, it is valuable to investigate the implications of these modelling decisions on benefit estimation. This research investigates the effects of alternative nesting structures, choice of estimator, and variable inclusion on welfare measurement.

Although recent work in recreation demand analysis and contingent valuation models have commonly presented interval estimates of welfare measures (Cameron (1991), Kling (1991), Park, Loomis, and Creel (1989), Duffield and Patterson (1989), among others), researchers employing discrete choice recreation demand models have not typically provided estimates of standard errors of the welfare measures. An additional contribution of this work is to describe and implement a simulation procedure adapted from work by Krinsky and Robb (1986 and 1991) to estimate standard errors for the welfare measures from nested logit models.

ISSUES IN ESTIMATING NESTED LOGIT MODELS AND COMPUTING WELFARE MEASURES

Suppose an individual is considering whether to go recreational fishing from the beach, pier, private boat or charter boat and that there are J_m sites at each of the four modes. There are numerous ways these alternatives could be grouped in estimating a nested logit model. One possibility is to group the sites together by mode. In this case, it is convenient to consider two decisions, choice of mode and choice of site. Again, a sequential interpretation is not necessary, all that is necessary is a hypothesis that there is correlation in the pattern of choices within modes.

Using this structure, the probability that an individual will choose to visit site j using mode m can be represented by the following

$$P_{jm}=P(j|m)P(m), \tag{1}$$

where P(j|m) is the probability that the individual chooses site j conditioned on choosing mode m and P(m) is the probability the individual will choose mode m.

Assuming a generalized extreme value distribution results in the nested logit model (Morey (1994); Maddala(1983)) where this probability can be expressed as

$$P_{jm}=\frac{e^{V_{jm}/\alpha_m}[\sum_{l=1}^{J_m} e^{V_{lm}/\alpha_m}]^{(\alpha_m-1)}}{\sum_{k=1}^{M}[\sum_{l=1}^{J_k} e^{V_{lk}/\alpha_k}]^{\alpha_k}}, \tag{2}$$

where V_{jm} is the utility from visiting site j using mode m, α_m is a parameter that measures the degree of substitution between the various modes, M is the number of modes, and J_m is the number of sites in mode m. Note that the number of sites may vary by mode. The coefficient, α_m, is referred to variously as the inclusive value coefficient or the dissimilarity parameter. When $\alpha_m=1$, for all m, the probability expression in (2) collapses to the standard multinomial logit probability, where the IIA property holds between all alternatives. Thus, one test of the IIA hypothesis is whether $\alpha_m=1$.

Applications of the nested logit model in the recreation demand literature have generally estimated a single value for all α_m, i.e., $\alpha_1=\alpha_2=...=\alpha$. The indirect utility function is typically specified as linear; a simple example is

$$V_{jm}=\beta_1 P_{jm}+\beta_{2m} C_{jm}, \tag{3}$$

where greek letters are coefficients, P_{jm} and C_{jm} are price and catch variables that describe the utility associated with visiting site j using mode m. Note that the coefficient on catch is subscripted by mode (m) to denote the possibility that the coefficient value might vary across modes.

FIML estimation is accomplished by defining the log likelihood function to be the product over a sample of individuals of the probability statements in (1).

$$L = \prod_{i=1}^{N} P_{ijm}, \tag{4}$$

where the i subscripts individuals in the sample and there are a total of N observations. (For ease of exposition, the i subscript will be omitted). Morey, Rowe, and Watson (1993) employ a FIML estimator in their study of recreational angling in Maine.

An alternative formulation of this choice probability that makes the two part decision structure and the sequential estimation strategy explicit is

$$P(j|m) = \frac{\exp(V_{jm}/\alpha_m)}{\exp(I_m)},$$

$$P(m) = \frac{\exp(\alpha_m I_m)}{\sum_{k=1}^{M} \exp(\alpha_k I_k)}, \tag{5}$$

$$I_m = \log(\sum_{i=1}^{J_m} \exp(V_{im}/\alpha_m)).$$

Here, I_m is referred to as the inclusive value and is a measure of the expected maximum utility from the alternatives contained in that subgroup. Sequential estimation proceeds by first estimating the conditional choice probability $P(j|m)$. This provides estimates of the coefficient vector β/α which are used to construct the inclusive values. These values are then used in a second stage to estimate α from which the individual β's can be recovered.

As Brownstone and Small (1989) note, when employing the sequential estimation strategy it becomes difficult to allow the α's to vary by mode (m) since the values of the β's are recovered by dividing by the inclusive value coefficient (α). Thus, the equations in (4) are commonly written with an α that is uniform across modes. Morey (1994) notes that common usage of the term "nested logit" carries with it the assumption of α's that do not vary across modes. Most applications of the nested logit model in recreation demand analysis have employed the sequential estimator.

As noted earlier, problems identified with the use of sequential estimation include asymptotic inefficiency and inconsistent standard error estimates. Also, FIML estimates of the coefficients may differ substantially from sequential estimates in finite samples.

Welfare measures computing the value of angling per choice occasion can be constructed based on estimates from either estimation strategy. Hanemann demonstrated that the compensating variation associated with a change in prices or qualities can be expressed as

$$cv = -\frac{1}{\beta_1}\{\ln[\sum_{m=1}^{M} \alpha_m[\sum_{j=1}^{J_m} e^{\alpha_m V^2_{mj}}]^{(1/\alpha_m)}] - \ln[\sum_{m=1}^{M} \alpha_m[\sum_{j=1}^{J_m} e^{\alpha_m V^1_{mj}}]^{(1/\alpha_m)}]\} \qquad (6)$$

where the superscripts on V indicate whether the prices and qualities are set at the new level (superscript 2) or the original level (superscript 1). These welfare measures will be compared to measures resulting from simple multinomial logit models with no nesting as well as to nested models estimated simultaneously.

Since the estimated coefficients in equation (6) are random variables, the compensating variation estimate is also random. A simple procedure for estimating the standard error of the welfare measure in (6) is to follow the suggestion of Krinsky and Robb (1986) where they employed a simulation strategy to estimate standard errors for elasticities.

In the context of welfare measurement, the procedure is as follows. First, upon estimating the model, denote the estimated coefficient vector β and its estimated variance-covariance matrix, $\hat{VC}$. Now, take a random draw from the multivariate normal distribution with a mean of β and variance-covariance matrix $\hat{VC}$. This provides a pseudo parameter vector that can be denoted, β^*. Plugging these values into equation (6) yields a pseudo-welfare estimate, cv^*. By repeating this procedure a large number of times, a distribution for cv is constructed. The standard error of this distribution is an estimate of the standard error of cv.

In addition to estimating the standard error, this procedure can also be used to estimate bias. An estimate of the bias is provided by the difference between the average cv^* and the point estimate of cv. If this difference is positive, it suggests that the point estimate is too large, if the difference is negative, it suggests downward bias. These bias estimates can be used to adjust the point estimates of welfare ex post.

This simulation procedure will be applied to several models estimated using the FIML estimator. Note that it would be difficult to apply the procedure to estimates from the sequential model since a single variance-covariance matrix is not estimated. In particular, the covariance terms between the dissimilarity coefficient (α) and the β's are not directly estimated. The relative ease of computing standard errors using the Krinsky and Robb procedure from a FIML estimator is another advantage of FIML over sequential estimation.

DATA

The data used for this study are from the Southern California Sportfish Economic Survey sponsored by the National Marine Fisheries Service, Southwest Fisheries Science Center and the California Department of Fish and Game and was conducted in 1989 (Thomson and Crooke, 1991). The survey collected information from a random telephone survey of households in the eight coastal counties of southern california in four waves during 1989. A follow-up mail survey collected detailed information from anglers on their most recent fishing trip. Noncoastal county anglers were contacted via the Marine Recreational Fishery Statistics Survey and sent the mail survey if they indicated willingness to participate.

In their responses to the mail survey, anglers reported their zip code, which site they visited on their most recent saltwater fishing trip, and which of four modes of fishing they chose: pier, beach, private boat, or charter boat. They were provided with a map of southern California that identified 34 fishing locations along the seven southern California coastal counties. The area covered ranges from the Mexican border north to San Simeon. The map identified 22 coastal sites where shore or offshore fishing could take place and 12 offshore sites that could be reached only by boat.

Since logit models require information for each observation on each alternative, and since individuals only reported distances and travel costs for the location they actually visited, the distances between their origin zip code and each of the shore sites was computed. Roundtrip travel costs for each individual to each site were computed by multiplying the roundtrip distances by a constant cost per mile and adding an opportunity cost of travel time. The cost of time was constructed simply as the after tax wage rate times an estimate of the time spent travelling to the site.[2] Finally, boat fees and fuel costs were added to the charter and private boat alternatives.

Exogenous data on catch rates were obtained from the 1989 Marine Recreational Fishery Statistics Survey. This survey intercepted anglers on site and collected data on catch rates. This data contained information on average catch per hour fished of each of the major species in the area for each of the four modes. This data was only available for more aggregate areas than those corresponding to the mail survey. Since the MRFSS data was the lowest common denominator, the site identified by it were employed in this study. Once this data was combined with the survey data there were a total of 5 sites where anglers could fish from the beach or pier and an additional 3 sites that could be accessed by party or charter boat. Thus, there were a total of 26 mode/site alternatives: 5 pier sites, 5 beach sites, 8 private boat sites, and 8 charter boat sites.

In the mail survey, anglers reported which species they were primarily fishing for; in many cases, they indicated they were fishing for two or more species on the same trip. To more accurately capture the catch rates of interest, a catch rate variable was constructed for each individual that was the sum of the catch rates of each of the species they identified as having been primarily pursuing. This construction appears to more accurately depict the desirability of a site from the perspective of an angler with a particular species in mind.

In addition to prices and catch, two other variables were used in some estimations to help predict angler's site selection. Charter boating is likely to require additional effort and organization than the three other modes. Often reservations must be made in advance and groups may be organized to go as a party. Additionally, charter boating may have significant additional costs that are not reflected in boat fees or travel costs. For these reasons, a dummy variable taking on the value of "1" when the choice was charter boat and "0" otherwise was constructed. A second dummy variable identifying whether the angler was

[2] In a recent extension of Bockstael, Hanemann, and Strand's (1987) work on the value of time, Larson (1993) argues that there are many cases when valuing time at the full after-tax wage rate is applicable.

targeting at least one species was also constructed. This variable took on a value of "1" if the angler was targeting a species and chose one of the offshore modes: private boat or charter boat.

EMPIRICAL SPECIFICATIONS AND ESTIMATES

In the typical recreation demand application of the nested logit model, the researcher(s) hypothesize a particular tree structure, choose variables to be included in the estimation, and estimate the model sequentially. The resulting coefficient estimates are used to estimate welfare associated with hypothesized changes in catch rates or site availability. The purpose of the empirical exercise reported here is to examine the sensitivity of welfare estimates to variations in the set of maintained hypotheses. Specifically, four alternative nesting structures are employed to estimate nested logit models. Three different subsets of variables are used to describe the choices individuals make between different sites and modes. The models are estimated using the sequential estimator and FIML and they are estimated allowing the dissimilarity parameters to vary by mode and the coefficient on catch rates to vary by mode as well. Finally, the results of these model estimations are used to compute point estimates of welfare for closing down all of the beach, pier, private and charter boat sites. Standard errors for welfare measures computed via Krinsky and Robb's simulation method are reported for the four nesting structures.

MNL Estimation

As a starting point, a basic multinomial logit model is estimated implying that IIA holds between all alternative site/mode combinations. This is equivalent to estimating the model identified by (2), (3), and (4) assuming the $\alpha_m=1$, for all m. In this simple model, there are a total of 26 alternative site/mode combinations, thus 26 alternatives.[3]

Table 1 contains the results of the estimation when just two variables are included: price and catch rate. As for all of the model estimates presented here, the coefficient on price is negative and significant.

[3]The model was estimated using both the discrete choice option in LIMDEP and maximizing the log likelihood function explicitly in TSP. The coefficient estimates always agreed to within at least two decimal places.

Likewise, the catch rate variable is also significant. Since the focus of this paper is on implications of model specification for welfare estimation, individual coefficients are not reported for each of the models. Rather, welfare measures and likelihood values are reported. In addition, when the models exhibit either unexpected sign values of insignificant coefficients, these will be noted in the discussion.

The first column of table 1 contains six welfare estimates and the value of the log likelihood function for the multinomial logit (MNL) model. The first four welfare estimates are the average willingness to pay per choice occasion if the 5 beach alternatives, 5 pier alternatives, 8 charter boat alternatives, and 8 private boat alternatives are shut down, respectively. The offshore measure is the welfare estimate corresponding to elimination of all pier and beach fishing alternatives. The shore measure is the estimate associated with eliminating all charter boat and private boat fishing alternatives. As noted in the Jones and Stokes report (1987), the sum of the welfare measures from shutting down the separate pier and beach alternatives is less than the shore measure and likewise for the offshore estimate relative to charter and private boat modes. Thus, there are increasing costs to anglers of closing more sites.[4]

The MNL estimates indicate that the average willingness to pay per angler per choice occasion to maintain access to beach sites is about \$8.75. The value of pier alternatives is about the same at \$8.15. As expected, charter and private boat alternatives are more valuable then pier and beach alternatives. Also, offshore alternatives are more valuable than shore alternatives.

A second MNL model was estimated allowing the coefficients on catch rate to vary across the four modes. The welfare estimates resulting from this estimation are reported in column five. The coefficient on price was not allowed to vary across modes as there are theoretical arguments for maintaining uniformity in the cost coefficient across alternatives (Truong and Hensher 1985)). Additionally, the two dummy variables do not vary across modes so allowing their coefficients to vary by mode is not possible. Each of the coefficients on catch are positively signed and significant. A likelihood ratio test suggests that the MNL with differences in the catch coefficient provides a significantly better fit. The welfare estimates are

[4] Note that one implication of the logit model as specified here is that angling is a necessary good. If all of the site/mode combinations are eliminated, utility for the angler goes to zero.

not hugely different from the MNL model with uniform catch coefficients, but clear differences are apparent.

Both of these models were estimated again using price, catch, and the target dummy (table 2) and price, catch, the target dummy, and the charter boat dummy (table 3) as variables. The same welfare measures were estimated and reported. Not surprisingly, there are differences in the welfare estimates across the three sets of variables. Likelihood ratio tests suggest that the inclusion of all four variables is appropriate, but all three tables are reported here to examine some interesting comparisons that arise.

FIML vs Sequential Estimation

To allow for violations of the IIA assumption, nested multinomial logit models are estimated. The first nesting structure examined is one where the sites are grouped together by mode. Nesting structure A in Figure 1 identifies the structure where the five beach alternatives are grouped together, as are the five pier alternatives, the eight charter alternatives and the eight private alternatives.

The first estimation strategy is sequential, corresponding to the probability statements expressed in equation (5). Estimation proceeds in reverse order. First the choice among sites is estimated. Based on coefficient estimates from this estimation, the inclusive value is computed and used to estimate the choice among modes.

When adding the dummy variables, charter and target, another disadvantage of sequential estimation appears. Since these dummy variables do not vary across sites within a mode, they cannot enter the first stage of the estimation. Rather, they must be introduced in the second stage. When estimating the models sequentially, these two variables enter the second estimation along with the inclusive value computed from the first stage.

Results for the sequential estimator are presented in the second column of tables 1-3. In addition to the six welfare estimates, the coefficient on the dissimilarity parameters are presented in the 8th row. Standard errors for the dissimilarity parameters are presented in parentheses next to the coefficient estimates.

The three estimated inclusive value coefficients are 0.22, 0.29, and 0.34. All are statistically significant at the 1% level. They are also significantly different from 1.0, indicating rejection of the IIA hypothesis.

The third column in the three tables presents results from FIML estimation.[5] All of the coefficients are significant and have the anticipated sign. The dissimilarity parameters are 0.28, 0.40, and 2.43. The dissimilarity parameter estimated from the sequential model and FIML in table 1 are fairly close (0.22 vs 0.28). The corresponding welfare measures differ, but are arguably quite similar. When an additional variable is included (table 2) the dissimilarity parameters diverge a bit more and when the second dummy variables is added (table 3) they diverge substantially (2.43 vs. 0.34). This result is likely due to the differential way of including the dummy variables into the estimation. Using FIML, the variables enter the model in the same way that parameters that vary across modes does. Using the sequential estimator, they can enter only in the second stage and thus contribute less information to the choice process. Note the substantial differences in welfare estimates from the sequential vs. FIML estimator in table 3.

Additionally, the other coefficient estimates also differ considerably from the sequential estimators yielding quite different welfare estimates as can be seen by comparing columns 2 and 3 across the three tables. These quite disparate results add further to the results of Cameron, Hensher, and Brownstone and Small that suggest that the sequential estimator may yield quite different coefficient estimates from FIML.

A further advantage of the FIML estimator is the ease with which nonuniformity in parameter values can be estimated. Each of the three models was re-estimated with additional flexibility in the parameter values. First, the inclusive value coefficients were allowed to vary across modes. This is identified in the tables as FIML ($\alpha\neq$). Next, the coefficient on catch were allowed to vary by mode while imposing a uniform inclusive value (labelled FIML ($\beta_2\neq$)). Finally, both the inclusive value and the coefficient on catch were allowed to be mode specific (FIML ($\alpha\neq,\beta_2\neq$)). In these models, the coefficients on all of the variables except pier and beach catch rates are of the expected sign and significant. For the pier and beach catch rates, the signs are occasionally negative and often insignificant.

[5]Estimation was performed by specifying the likelihood function in TSP and using the ML routine.

Eliminating the restrictions of uniformity of parameters across modes can make dramatic differences in the welfare estimates resulting from the models, as seen in the tables. For example, using the FIML estimator with three variables (table 2), one can get welfare estimates ranging anywhere from \$2.32 to \$22.12 for elimination of beach access, depending on whether the α's and β_2's are allowed to vary by mode or not. Equivalently large differences are evident throughout the three tables.

A perusal of the equivalently large differences in estimated dissimilarity coefficients is also instructive. When more than one dissimilarity value is estimated, the far left column of the table identifies which mode is associated with that coefficient. Beginning with table 3, all of the dissimilarity coefficients estimated via FIML are greater than unity and substantially larger than that estimated using the sequential estimator (0.34). Also, they are all significantly different from both 0 and 1. In table 2, the dissimilarity coefficients do not diverge as much between the sequential and FIML estimators when the coefficient is equal across modes, but are quite large when they are not constrained to be equal across modes. A similar pattern appears in table 1.

Not surprisingly, the use of the dissimilarity coefficient to test for the IIA property depends critically upon model structure and choice of variables. In all cases in tables 1-3, the IIA property would be rejected based on statistical significance of the inclusive value coefficients and likelihood ratio tests. However, the magnitude of the value varies from 0.22 to as high as 2.69.

An intuitive interpretation of the dissimilarity parameters has to do with the degree of substitution between alternatives within a group and between groups. Train, McFadden, and Ben-Akiva (1987) argue that dissimilarity parameters greater than on imply that there is more substitution between groups than within and dissimilarity parameters less than on imply there is more substitution with groups than between. Naturally, values equal to on indicate an equal degree of substitutability between all alternatives.

McFadden (1981) identified that a globally sufficient condition for the nested logit model to be consistent with utility maximization is that $0<\alpha<1$. As noted above, many of the estimated values do not conform to these requirements. This pattern has been noted in other studies. Recently, Borsch-Supan (1985 and 1990) has demonstrated local sufficiency conditions that permit values of $\alpha>1$.[6] His work establishes that dissimilarity coefficients that exceed one can be consistent with utility theory.

Nesting Structure

To examine the implications of the grouping structure assumed by the researcher on welfare measures and coefficient estimates, three additional nesting structures are hypothesized and estimated. Figure 1 contains all four structures. These are, of course, just four of many possibilities that could be identified. In the first two structures (A and B), the groups are made by mode. In A, the alternatives in each of the four modes (pier, beach, private, and charter) are grouped together. In B, there are two modes: offshore (the combination of private and charter) and shore (the combination of pier and beach).

In model A, the sites are grouped together by the four 4 modes: beach, pier, private, or charter. In model B, the sites are grouped by just 2 aggregated modes: fishing from the shore (pier or beach) or fishing from offshore (private or charter). In model A, IIA is relaxed between choices from the four different groups, but maintained with each group. In model B, IIA is relaxed between alternatives in the shore and offshore group, but again not within the groups. Thus, in model B, an angler is assumed to view the 10 shore alternatives (5 pier/site choices + 5 beach/site choices) as perfect substitutes. Likewise, the 16 offshore alternatives are viewed as perfect substitutes.

Table 4 contains the welfare measures and dissimilarity parameters estimated using FIML for model B. All four variables are included in the estimation and α and β_2 are allowed to vary across the modes. All of the coefficients are significant and of the expected sign. The dissimilarity coefficients on the offshore mode is not significantly different from 1.00 suggesting that there is equal substitutability between shore and offshore alternatives as there are within offshore alternatives. The dissimilarity coefficient on the shore mode is significantly less than 1.00. The welfare values are similar to those presented for model A for elimination of beach and pier access, but quite a bit larger for the other four welfare measures.

[6] Unfortunately, given the large number of alternatives considered in this study (26) his conditions would require the evaluation of many derivatives.

An alternative approach to grouping angler's alternatives is by site. Again, two alternative forms of nesting structure are examined and identified as models C and D on Figure 1. In model C, four mode alternatives are grouped together by site. Note that for sites 1-5 there are 4 mode choices, but for sites 6-8 there are only 2 mode choices (charter and private boat). In model D, there are only two aggregate site alternatives, shore sites and offshore sites (i.e., sites 1-5 vs sites 6-8).

The welfare results for these two models are also presented in table 4. In model C, there are five insignificant coefficients (out of a total of 19 estimated parameters) and one negative catch coefficient. Model D has no insignificant coefficients. The dissimilarity parameters are generally significantly different from 1.00, with exceptions for the site 3 in model C and the shore sites in model D.

Again, differences in coefficient estimates yield large differences in welfare estimates. The measures estimated using model C are nearly twice as large as predicted by any of the other models.

ESTIMATING STANDARD ERRORS AND BIAS

Additional insight into the differences in welfare measures may be provided by the computation of standard errors of the estimates. Following the procedures described above for simulating the distribution of the welfare measure, standard errors were constructed for each of the welfare measures reported in table 4. These estimates are presented in parentheses below the point estimates.

For the most part, the welfare measures are estimated quite precisely. The exception occurs in model C where the standard errors are larger relative to the point estimate than for the other three models, but still fairly tight. The larger standard errors occur here due to the less precise estimate of the coefficients. As noted earlier, there were five coefficients that were not significant out of a total of nineteen. Higher standard errors on the coefficients translate into higher standard errors on the welfare measures.

The bias of the welfare estimators was also computed, but not reported in the table. In all cases the bias was quite low, ranging in absolute value from zero to just over three percent. Thus, though the point estimates could be adjusted to correct for the bias, its small magnitude here hardly makes it worthwhile.

FINAL COMMENTS

Sensitivity analysis performed on nested logit models have identified several issues of importance to researcher employing these models. First, sequential estimation has several disadvantages which FIML estimation does not. It was noted here that inclusion of variables that do not vary across all levels of the nest are more difficult with sequential models. Additionally, allowing variation in parameter values across levels of the nest is also much easier using a FIML estimator. Finally, and probably most importantly, it is much easier to construct standard error estimates for welfare measures coming from a FIML estimator due to the single variance-covariance matrix estimated.

The form of the nesting structure was identified as an important element in welfare measurement. Four alternative nesting structures were estimated and welfare estimates derived from them compared. The results from these estimations imply, as expected, that the welfare estimates and conclusions about substitutability between alternatives are inextricably tied to the maintained hypothesis about model structure. These maintained hypotheses can take many forms including decisions about which variables to group together in a nest, restrictions on parameter values across groups, and choice of estimator. Interestingly, however, in almost all cases, the IIA property is rejected in favor of correlation between groups of alternatives.

The use of the Krinsky and Robb simulation procedure to estimate standard errors of welfare measures is straightforward and can be applied to all models of nested models estimated simultaneously (using FIML). The additional effort to obtain standard errors to supplement the point estimates of welfare is relatively minimal.

The results from this study reinforce the findings of Smith, Desvousges, and Fisher (1986) who argue that judgment in model building is a crucial factor in the resulting benefit estimates. It is, of course, not a surprising finding, but one which applied researchers should be continually aware of. Sensitivity analysis and good judgment continue to be key ingredients in obtaining benefit estimates that are defensible and valuable enough for policy analyses.

Figure 1: Alternative Nesting Structures

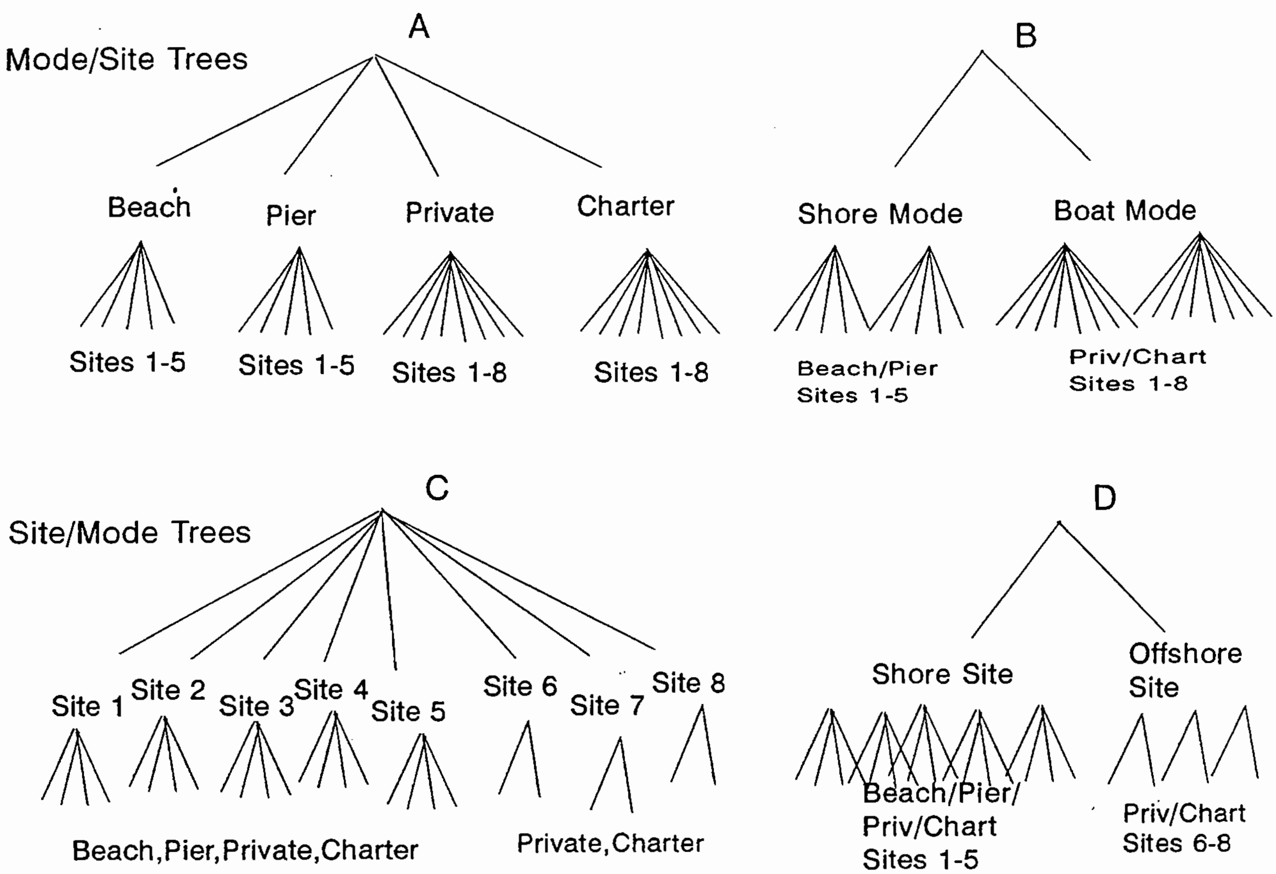


Table 1: Welfare Measures and Summary Statistics for Nested Logit Models: Four Levels in First Choice and Two Variables							
	MNL	SEQ	FIML	FIML ($\alpha\neq$)	MNL ($\beta_2\neq$)	FIML ($\beta_2\neq$)	FIML ($\alpha\neq,\beta_2\neq$)
Beach	8.74	35.20	29.23	3.71	6.53	14.83	3.62
Pier	8.16	34.47	28.44	3.79	7.92	16.98	4.26
Charter	11.10	36.51	31.65	10.59	11.31	21.41	11.59
Private	14.87	39.34	33.67	8.82	16.75	25.51	10.14
Offshore	32.85	92.31	79.88	27.46	36.95	59.77	33.38
Shore	19.67	83.20	68.62	8.40	16.40	36.91	8.89
α (private charter pier beach)	---	0.22 (0.07)	0.28 (0.07)	1.20(0.11) 1.94(0.17) 0.98(0.14) 0.88(0.11)	---	0.46 (0.07)	1.07(0.11) 1.77(0.17) 0.85(0.13) 0.84(0.11)
Log likelihood	3048	1390 1634	3023	2933	3031	3010	2926

Table 2: Welfare Measures and Summary Statistics for Nested Logit Models: Four Levels in First Choice and Three Variables							
	MNL	SEQ	FIML	FIML ($\alpha\neq$)	MNL ($\beta_2\neq$)	FIML ($\beta_2\neq$)	FIML ($\alpha\neq,\beta_2\neq$)
Beach	9.61	26.72	22.18	2.50	7.09	11.10	2.32
Pier	8.93	25.99	21.35	3.15	8.62	13.13	3.16
Charter	11.00	27.51	22.91	7.92	11.15	15.84	8.21
Private	13.84	30.27	24.61	6.49	16.26	20.20	7.20
Offshore	30.97	70.34	57.89	21.48	36.39	46.62	25.00
Shore	22.01	62.95	52.22	6.60	18.27	28.42	6.46
α (private charter pier beach)	---	0.29 (0.08)	0.40 (0.08)	1.70(0.12) 2.69(0.20) 1.42(0.16) 1.18(0.13)	---	0.63 (0.08)	1.57(0.13) 2.57(0.21) 1.36(0.17) 1.16(0.13)
Log likelihood	3032	1390 1631	3018	2876	3002	2995	2863

Table 3: Welfare Measures and Summary Statistics for Nested Logit Models: Four Levels in First Choice and Four Variables							
	MNL	SEQ	FIML	FIML ($\alpha\neq$)	MNL ($\beta_2\neq$)	FIML ($\beta_2\neq$)	FIML ($\alpha\neq,\beta_2\neq$)
Beach	6.01	21.40	1.90	2.05	4.95	1.96	2.17
Pier	5.82	20.72	1.85	2.66	5.84	2.33	2.97
Charter	13.97	28.50	5.75	7.05	14.19	6.72	8.08
Private	8.19	23.96	4.80	5.63	9.79	5.69	6.76
Offshore	29.22	64.87	15.92	19.16	33.61	18.96	24.19
Shore	13.62	49.66	4.35	5.48	12.36	4.91	6.36
α (private charter pier beach)	---	0.34 (0.08)	2.43 (0.17)	1.89(0.15) 2.59(0.22) 1.68(0.20) 1.42(0.17)	---	2.10 (0.18)	1.66(0.15) 2.25(0.22) 1.45(0.19) 1.24(0.15)
Log likelihood	2934	1390 1569	2903	2871	2900	2883	2857

Table 4: Welfare Measures and Standard Errors for Four Alternative Nesting Structures					
	MNL	Model A	Model B	Model C	Model D
Beach	6.01 (0.24)[1]	2.17 (0.31)	3.22 (0.25)	11.67 (2.30)	4.72 (0.19)
Pier	5.82 (0.23)	2.97 (0.38)	3.69 (0.32)	11.50 (2.26)	4.56 (0.19)
Charter	13.97 (0.72)	8.08 (0.90)	16.26 (1.02)	35.46 (6.56)	14.14 (0.76)
Private	8.19 (0.34)	6.76 (0.70)	13.70 (0.80)	27.90 (5.37)	9.79 (0.53)
Offshore	29.22 (0.95)	24.19 (2.26)	43.02 (3.92)	61.79 (11.20)	28.64 (1.37)
Shore	13.62 (0.62)	6.36 (1.14)	11.18 (1.39)	23.56 (3.68)	10.74 (0.51)
α	1.00 (not estimated)	1.66(0.15)[2] 2.25(0.22) 1.45(0.19) 1.24(0.15)	1.07(0.10)[3] 0.65(0.08)	1.94(0.52)[4] 2.21(0.54) 1.22(0.53) 2.33(0.56) 2.54(0.52) 2.95(0.96) 4.05(0.99) 4.21(1.01)	1.38(0.15)[5] 0.97(0.09)
Log likelihood	2934	2857	2860	2751	2845

[1] Standard error estimates computed using Krinsky and Robb's simulation method are reported below each welfare estimate.
[2] The four values for α correspond to the four modes: private, charter, pier, and beach.
[3] The two values for α correspond to the two modes: offshore angling and shore angling.
[4] The eight values for α correspond to the eight sites in ascending order.
[5] The two values for α correspond to two sets of sites: offshore sites and shore sites.

References

Bockstael, N.E., W.M. Hanemann, and C.L. Kling. "Measuring Recreational Demand in a Multiple Site Framework," **Water Resources Research** 23(1987): 951-960.

Bockstael, N.E., W.M. Hanemann, and I.E. Strand. "Time and the Recreation Demand Model." **American Journal of Agricultural Economics** 69(1987): 293-302.

Bockstael, N.E., K.E. McConnell, and I.E. Strand. "A Random Utility Model for Sportfishing: Some Preliminary Results for Florida," **Marine Resource Economics** 6(1989): 245-260.

Borsch-Supan, Axel. "Hierarchical Choice Models and Efficient Sampling; with Applications on the Demand for Housing." **Methods of Operations Research** (1985):175-187.

Borsch-Supan, Axel. **Econometric Analysis of Discrete Choice: With Applications on the Demand for Housing in the U.S. and West Germany,** Lecture Notes in Economics and Mathematical systems, Springer-Verlag, New York, 1987.

Borsch-Supan, Axel. "On the Compatibility of Nested Logit Models with Utility Maximization." **Journal of Econometrics** 43(1990): 373-388.

Brownstone D. and K.A. Small. "Efficient Estimation of Nested Logit Models, **Journal of Business and Economic Statistics,** 7(1989): 67-74.

Cameron, T.A. "A Nested Logit Model of Energy Conservation Activity by Owners of Existing Single Family Dwellings," **Review of Economics and Statistics** (1985): 205-211.

Cameron, T.A. "Interval Estimates of Non-Market Resource Values from Referendum Contingent Valuation Surveys," **Land Economics** 67(1991): 413-21.

Caulkins, P., R. Bishop, and N. Bouwes. "The Travel Cost Model for Lake Recreation: A Comparison of Two Methods for Incorporating Site Quality and Substitution Effects," **American Journal of Agricultural Economics** 68(1986): 291-297.

Creel, M. and J. Loomis. "Recreation Value of Water to Wetlands in the San Joaquin Valley: Linked Multinomial Logit and Count Data Trip Frequency Models," **Water Resources Research** 28(1992): 2597-2606.

Duffield, J. and D Patterson. "A Discrete Probability Model: Implications for Design and Interpretation of Logistic Contingent Valuation," In **Benefits and Costs in Natural Resources Planning,** Interim Report #2, Western Regional Research Project W-133 (1989).

Efron, B. "Nonparametric Standard Errors and Confidence Intervals," **Canadian Journal of Statistics** 82(198): 139-158.

Feenberg D., and E. Mills. **Measuring the Benefits of Water Pollution Abatement,** New York: Academic Press (1980).

Hanemann, W.M. "A Methodological and Empirical Study of the Recreation Benefits from Water Quality Improvement," Ph.D. dissertation, Department of Economics, Harvard University, 1978.

Hanemann, W.M. "Applied Welfare Analysis with Qualitative Response Models," California Agricultural Experiment Station, October, 1982.

Hensher, D.A. "Sequential and Full Information Maximum Likelihood Estimation of a Nested Logit Model," **Review of Economics and Statistics** 56(1986): 657-667.

Jones and Stokes Associates, "Southcentral Alaska Sport Fishing Economic Study," Prepared for Alaska Department of Fish and Game, November 1987.

Kling, C.L. "Estimating the Precision of Welfare Measures," **Journal of Environmental Economics and Management** 21(1991): 244-259.

Krinsky, I. and A. Robb. "On Approximating the Statistical Properties of Elasticities," **Review of Economics and Statistics** 68(1986):715-719.

Krinsky, I. and A. Robb. "Three Methods for Calculating the Statistical Properties of Elasticities: A Comparison," **Applied Economics** (1991).

Larson, D.M. "Separability and the Shadow Value of Leisure Time," **American Journal of Agricultural Economics** 75(1993):572-577.

Maddala, G.S. **Limited-Dependent and Qualitative Variables in Econometrics** Cambridge University Press, Cambridge (1983).

Milon, J.W. "A Nested Demand Shares Model of Artificial Marine Habitat Choice by Sport Anglers," **Marine Resource Economics** 5(1988): 191-213.

Morey, E. "Two RUMs UnCLOAKED: Nested-Logit Models of Site Choice and Nested-Logit Models of Participation and Site Choice," Discussion paper, Department of Economics, University of Colorado, Boulder, February 1994.

Morey, E.R., W.D. Shaw and R.D. Rowe. "A Discrete Choice Model of Recreational Participation, Site Choice, and Activity Valuation When Complete Trip Data Are Not Available," **Journal of Environmental Economics and Management** 20(1991).

Morey, E.R., R.D. Rowe, and M. Watson. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing," **American Journal of Agricultural Economics** 75(1993): 578-592.

Park, T., J. Loomis, and M. Creel. "Confidence Intervals for Evaluating Benefit Estimates from Dichotomous Choice Contingent Valuation Surveys," University of California, Davis, (1989).

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

Parsons, G.R. and M.S. Needelman. "Site Aggregation in a Random Utility Model of Recreation," **Land Economics** 68(1992): 418-433.

Smith, V.K., W.H. Desvousges, and A. Fisher. "A Comparison of Direct and Indirect Methods for Estimating Environmental Benefits," **American Journal of Agricultural Economics** 68(1986): 280-290.

Thomson, C.J. and S.T. Crooke. "Results of the Southern California Sportfish Economic Survey," NOAA Technical Memorandum, National Marine Fisheries Service, Southwest Fisheries Center, August, 1991.

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

Truong, P. and D.A. Hensher. "Measurement of Travel Time Values and Opportunity Cost from a Discrete Choice Model," **Economic Journal** 95 (1985): 438-451.

Weak Complementarity and Other Normative Speculations

by

Jeffrey T. LaFrance*

Department of Agricultural & Resource Economics

The University of Arizona

Tucson, Arizona 85721

Abstract

This paper compares welfare measurement under weak complementarity with an uncountable number of alternative hypotheses that can be invoked to derive the welfare effects of a nonmarket good from measurements on market choices. All of the alternatives generate cardinal transformations of preferences but do not generate any testable implications for observable market choices. This indeterminacy is a particular problem in the context of nonmarket valuation because the nature of the particular transformation chosen dramatically affects the magnitude, and in some cases the sign, of the welfare measures obtained. The normative nature of weak complementarity and weak neutrality assumptions are illustrated numerically for the common case of a linear market demand function for the market good of interest.

*The helpful comments of Kenneth Arrow, Bruce Beattie, Bonnie Colby, Dennis Cory, Erwin Diewert, Rob Innis, Dale Jorgenson, and V. Kerry Smith are gratefully acknowledged. Remaining errors are, of course, my responsibility.

Weak Complementarity and Other Normative Speculations

1. INTRODUCTION

The focus of this paper is a common practice in applied welfare analysis - using measurements on the demands for a group of market goods to generate exact welfare measures for changes in the level of provision of a nonmarket good. The most common approach to evaluating the benefits from environmental improvement is to assume that a nonmarket quality is ***weakly complementary*** (Mäller) to a set of purchased market goods (Bockstael, Hanemann, and Kling; Bockstael and Kling; Bockstael and McConnell; Braden and Kolstad; Freeman; Hanemann; Larson 1991, 1992b; Mäller; McConnell; Mendelsohn 1984; Smith and Desvousges; Smith, Desvousges, and McGivney). Very recently, Larson and Flacco extended the weak complementarity methodology to the measurement of option values and option prices (Arrow and Fisher; Cicchetti and Freeman; Schmalensee). An alternative hypothesis to weak complementarity, called ***weak neutrality***, was recently proposed by Neill and analyzed in detail by Larson (1992a).[1] Larson (1992a, 1992b) argues that, in contrast to weak complementarity, wherein the inherent non-use value is zero by construction, this latter hypothesis permits the measurement of non-use values with market data. However, a great deal of disagreement continues over this issue (Cameron 1992a, 1992b; Chicchetti and Wilde; McConnell; Mendelsohn 1987; Smith).

In this paper, I show that ***all*** assumptions like weak complementarity generate ***cardinal*** transformations of consumer preferences. Because no testable implications for the observable choices of market goods result from these transformations (LaFrance and Hanemann), the welfare measurements obtained are

---

[1] These two hypotheses are similar in nature and in the degree of restriction imposed on preferences. Weak complementarity requires that, for some reference price vector, the price-weighted sum of all compensated demand responses to a change in the nonmarket quality vanishes. Weak neutrality requires that, at some reference point for market prices, at least one compensated quality response vanishes (Larson 1992a).

arbitrary and can not be defended against the widely divergent numbers that arise from an uncountable number of alternative, equally plausible, and equally indefensible assumptions.

These issues are illustrated with a linear demand for one market good. In the illustration, I show that the weak complementarity transformation affects the size of welfare estimates most in the range of values for the relevant demand parameters that are common in empirical work on nonmarket valuation. These results strongly suggest that the automatic invocation of weak complementarity and similar assumptions warrants serious reconsideration.

2. WEAK COMPLEMENTARITY AND WELFARE MEASUREMENT

Let $\boldsymbol{x} \in \boldsymbol{X} \subset \boldsymbol{R}_+^{\mathrm{n}}$ be the vector of consumption levels for the market goods of interest with $\boldsymbol{p} \in \boldsymbol{P} \subset \boldsymbol{R}_+^{\mathrm{n}}$ the corresponding price vector; let $z \in \boldsymbol{Z} \subset \boldsymbol{R}_+^{\mathrm{m}}$ be the vector of all other goods with $\boldsymbol{q} \in \boldsymbol{Q} \subset \boldsymbol{R}_+^{\mathrm{m}}$ the corresponding price vector; let $\tau \in \boldsymbol{T} \subset \boldsymbol{R}$ denote the nonmarket quality; and let income be $y \in \boldsymbol{Y} \subset \boldsymbol{R}_+$. We assume throughout that the utility function, u: $\boldsymbol{X} \times \boldsymbol{Z} \times \boldsymbol{T} \rightarrow \boldsymbol{U} \subset \boldsymbol{R}$, is twice continuously differentiable in (x,z,τ) and strictly increasing and strictly quasiconcave in (x,z).

To simplify the discussion, we consider a single Marshallian ordinary demand[2] function for the good, x, which depends on its price, p, the prices of other goods, $\boldsymbol{q}$, income, y, and a nonmarket quality, τ,

(1) $$x = h(p,\boldsymbol{q},y,\tau).$$

The Hicksian (compensated) demand for x satisfies Hotelling's lemma

(2) $$x = g(p,\boldsymbol{q},u,\tau) \equiv h(p,\boldsymbol{q},e(p,\boldsymbol{q},u,\tau),\tau) \equiv \partial e(p,\boldsymbol{q},u,\tau) / \partial p,$$

where $e(p,\boldsymbol{q},u,\tau)$ is the ***expenditure function*** defined by

(3) $$e(p,\boldsymbol{q},u,\tau) \equiv \min\{px + \boldsymbol{q}'z\text{: } u(x,z,\tau) \geq u,\ (x,z) \in \boldsymbol{X} \times \boldsymbol{Z}\}.$$

The solution to the partial differential equation defined by the two expressions in the middle and right-hand-side of (2) has the form

---

[2] It is straightforward to generalize the arguments that follow to groups of market goods following the line of reasoning developed by Bockstael and Kling.

$$(4) \qquad e(p,q,u,\tau) = \varepsilon(p,q,\tau,\theta(q,u,\tau)),$$

where $\theta(q,u,\tau)$ is a constant of integration that depends on (q,u,τ) but is not a function of p and $\varepsilon(p,q,\tau,\theta)$ is called the ***quasi-expenditure function*** (Hausman).

Inverting $\varepsilon(p,q,\tau,\theta)$ with respect to θ and then $\theta(q,u,\tau)$ with respect to u gives the indirect utility function in the form

$$(5) \qquad v(p,q,y,\tau) = \psi(q,\tau,\varphi(p,q,y,\tau)),$$

where $\varphi(p,q,y,\tau) = \theta(q,u,\tau)$ is called the ***quasi-indirect utility function*** (Hausman).

The ***weak complementarity hypothesis*** (Mäller), hereafter denoted by WCH, can be expressed in terms of the expenditure function as follows. Let $\partial e(p,q,u,\tau)/\partial p = 0$ implicitly define the ***choke price***, $p^\circ(q,u,\tau)$, for the compensated demand for x. Then x is weakly complementary to τ if and only if

$$(6) \qquad 0 \equiv \frac{\partial e(p^\circ(q,\mathrm{u},\tau),q,u,\tau)}{\partial\tau}$$

$$\equiv \frac{\partial\varepsilon(p^\circ(q,u,\tau),\tau,\theta(q,u,\tau))}{\partial\tau} + \frac{\partial\varepsilon(p^\circ(q,u,\tau),\tau,\theta(q,u,\tau))}{\partial\theta}\cdot\frac{\partial\theta(q,u,\tau)}{\partial\tau},$$

since $\partial\varepsilon(p^\circ(q,u,\tau),\tau,\theta(q,u,\tau))/\partial p \equiv 0$ by the definition of p° (Willig; Larson 1991). This defines a partial differential equation whose solution is of the general form

$$(7) \qquad \tilde{e}(p,q,u,\tau) = \tilde{\varepsilon}(p,q,\tau,\tilde{\theta}(q,u)),$$

and $\tilde{\varepsilon}(p,q,\tau,\tilde{\theta}(q,u))$ satisfies $\partial\tilde{\varepsilon}(p^\circ(q,u,\tau),q,\tau,\theta(q,u))/\partial\tau \equiv 0$. Inverting $\tilde{\varepsilon}(p,q,\tau,\tilde{\theta}(q,u))$ with respect to $\tilde{\theta}(q,u)$ and then inverting $\tilde{\theta}(q,u)$ with respect to u gives the transformed WCH indirect utility function in the form

$$(8) \qquad \tilde{v}(p,q,y,\tau) = \tilde{\psi}(q,\tilde{\varphi}(p,q,y,\tau)).$$

The difference between $\tilde{\psi}(q,\tilde{\varphi}(p,q,y,\tau))$ and $\psi(q,\tau,\varphi(p,q,y,\tau))$ from the perspective of measuring the welfare effects of changes in τ is that $\tilde{\varphi}(p,q,y,\tau)$ relays all of the information contained in $\tilde{\psi}(q,\tilde{\varphi}(p,q,y,\tau))$ relating to τ, while $\varphi(p,q,y,\tau)$ only reflects a subset of the information contained in $\psi(q,\tau,\varphi(p,q,y,\tau))$ that relates to τ. But ***any*** transformation that accomplishes this would work equally well without affecting the observed demands for the market good x. This is why assumptions like the WCH are ***not testable*** (i.e., neither empirically verifiable nor refutable), as recently shown by LaFrance and

Hanemann. This is no cause for concern when we are interested in the welfare effects of changes in the price of x or in income, because the exact welfare measures for such changes are invariant to ***all*** specifications for $\theta(q, \mathrm{u}, \tau)$ that are consistent with the utility maximization hypothesis (LaFrance and Hanemann). The difficulty with the WCH (and its relatives) is that for the very reason that this hypothesis ***permits*** welfare measurement with observations on market choices, it also introduces a ***cardinal*** transformation of preferences. The cardinal nature of this transformation influences the size, and at times even the sign, of the welfare measures obtained.

The nature of these issues are clear when we consider a specific alternative normative hypothesis (ANH) that suggests itself as an obvious choice, namely,

(9) $$\partial\theta(q, \mathrm{u}, \tau) / \partial\tau \equiv 0.$$

That is, the welfare implications of changes in τ are completely and directly recovered from the demand for x without any further transformation of the quasi-indirect utility function. This ANH is, in a sense, more convenient (and perhaps more natural) than the WCH since it represents the simplest, most primitive form of indirect preferences consistent with observed market choices. Denoting the solution to (9) by $\hat{\theta}(q,u)$, the resulting expenditure function has the form

(10) $$\hat{e}(p,q,u,\tau) = \varepsilon(p,q,\tau,\hat{\theta}(q,u)),$$

where $\varepsilon(\cdot)$ has the same structure in (10) as in (4) - only the properties of the constant of integration have changed.

From the standpoint of calculating the welfare effects of a change in τ, (7) and (10) have the same structure, since $\partial\tilde{\theta}(q,u)/\partial\tau \equiv \partial\hat{\theta}(q,u)/\partial\tau \equiv 0$. In other words, the ANH and the WCH appear to be equivalent in this regard. However, this superficial congruence is misleading because the structure of $\varepsilon(p,q,\tau,\hat{\theta})$ with respect to τ often is quite different from the structure of $\tilde{\varepsilon}(p,q,\tau,\tilde{\theta})$ with respect to τ due to the transformation of the quasi-expenditure function introduced to obtain the WCH. This difference can and often does lead to substantially different welfare measures in practice. Furthermore, without loss in generality,

$$\frac{\partial \tilde{e}(p,q,u,\tau)}{\partial u} \equiv \frac{\partial \tilde{\varepsilon}(p,q,\tau,\tilde{\theta}(q,u))}{\partial \tilde{\theta}} \cdot \frac{\partial \tilde{\theta}(q,u)}{\partial u} > 0,$$

implies $\partial \tilde{\varepsilon}(p,q,\tau,\tilde{\theta}(q,u)) / \partial \tilde{\theta} > 0$, while

$$\frac{\partial \hat{e}(p,q,u,\tau)}{\partial u} \equiv \frac{\partial \varepsilon(p,q,\tau,\hat{\theta}(q,u))}{\partial \hat{\theta}} \cdot \frac{\partial \hat{\theta}(q,u)}{\partial u} > 0,$$

implies $\partial \varepsilon(p,q,\tau,\hat{\theta}(q,u)) / \partial \hat{\theta} > 0$. Therefore, differences in the structure of the quasi-expenditure function with respect to $\tilde{\theta}(q,u)$ relative to $\hat{\theta}(q,u)$ also translate into differences in welfare measures. We conclude that, in general, the ANH and the WCH are not equivalent.[3] Furthermore, even if this particular ANH and the WCH are equivalent, an uncountable host of other ANH-like possibilities exist. For example, we could assert that $\theta(q,u,\tau) = \bar{\theta}(q,u)\tau^{\sigma}$ for some constant σ. ***No amount of data on the choices for x could refute or verify this claim.*** Moreover, arbitrary choices regarding the value of σ determine the sign and size of the welfare "estimates" for the value of a change in τ.

At this point, it is worthwhile to present a formal argument in the case of compensating variation. Consider a change in the nonmarket quality from τ° to τ', holding *(p,q,y)* fixed. Without imposing either the WCH or the ANH, the compensating variation for this change, c, is defined by

(11) $$u^{\circ} = \psi(q,\tau^{\circ},\varphi(p,q,y,\tau^{\circ})) = \psi(q,\tau',\varphi(p,q,y-c,\tau')),$$

while the compensating variation that results from the WCH, $\tilde{c}$, is defined by

(12) $$u^{\circ} = \tilde{\psi}(q,\tilde{\varphi}(p,q,y,\tau^{\circ})) = \tilde{\psi}(q,\tilde{\varphi}(p,q,y-\tilde{c},\tau')),$$

and the compensating variation for the ANH, $\hat{c}$, is defined by

(13) $$u^{\circ} = \hat{\psi}(q,\varphi(p,q,y,\tau^{\circ})) = \hat{\psi}(q,\varphi(p,q,y-\hat{c},\tau')),$$

where $\varphi(p,q,y,\tau)$ has the same structure in (11) as in (13).

---

[3] The two hypotheses are equivalent if and only if $\partial \varepsilon(p^{\circ}(q,u,\tau),\tau,\theta(q,u,\tau)) / \partial \tau \equiv 0$, as in the case of a price-elastic log-log demand model.

All that we know about the structure of ψ with respect to τ is that ψ is increasing in φ and the structure of φ is recovered from the demand for x. But since we can not determine the remaining structure of ψ with respect to τ, c cannot be calculated from (11). However, $\tilde{c}$ and $\hat{c}$ can be derived from $\tilde{\varphi}$ and φ, respectively. This follows from the fact that $\tilde{\psi}$ and $\hat{\psi}$ are monotonic in $\tilde{\varphi}$ and φ, respectively.

From duality theory, $\tilde{\varphi}(p,\boldsymbol{q},y,\tau) \equiv \varphi(p,\boldsymbol{q},y,\tau)$ if and only if $\tilde{\varepsilon}(p,\boldsymbol{q},\tau,\theta) \equiv \varepsilon(p,\boldsymbol{q},\tau,\theta)$ for all fixed values of θ in the domain of definition for either $\tilde{\varepsilon}$ or ε. In general, however, precisely because a transformation of the primitive quasi-expenditure function is necessary to satisfy the WCH, the latter condition is not satisfied. Therefore, it follows that $\tilde{c} \neq \hat{c}$, except perhaps in special circumstances. As the next section illustrates, the required circumstances usually are not met in environmental valuation studies.

3. AN ILLUSTRATION WITH A LINEAR DEMAND MODEL

Let the ordinary demand for x be given by

$$x = \alpha(\boldsymbol{q}) + \beta\left(\frac{p}{\pi(\boldsymbol{q})}\right) + \gamma\left(\frac{y}{\pi(\boldsymbol{q})}\right) + \delta\tau, \tag{14}$$

where $\alpha(\boldsymbol{q})$ is homogeneous of degree zero and $\pi(\boldsymbol{q})$ is linearly homogeneous in $\boldsymbol{q}$. We assume throughout this section that $\beta < 0$ and $\gamma > 0$, so that x is a normal good with a downward-sloping demand curve. The expenditure function for this demand equation has the form

$$e(p,\boldsymbol{q},u,\tau) = \theta(\boldsymbol{q},u,\tau)\cdot\exp\left\{\gamma\left(\frac{p}{\pi(\boldsymbol{q})}\right)\right\} - \left(\frac{\pi(\boldsymbol{q})}{\gamma}\right)\cdot\left[\alpha(\boldsymbol{q}) + \beta\left(\frac{p}{\pi(\boldsymbol{q})}\right) + \delta\tau + \left(\frac{\beta}{\gamma}\right)\right], \tag{15}$$

where $\theta(\boldsymbol{q},u,\tau)$ is linearly homogeneous in $\boldsymbol{q}$, increasing in u, and has a completely unknown structure with respect to τ. The quasi-indirect utility function is found by setting $y = e(p,\boldsymbol{q},\mathrm{u},\tau)$ and solving for θ,

$$\varphi(p,\boldsymbol{q},y,\tau) = \left(\frac{1}{\gamma}\right)\cdot\left[\alpha(\boldsymbol{q}) + \beta\left(\frac{p}{\pi(\boldsymbol{q})}\right) + \gamma\left(\frac{y}{\pi(\boldsymbol{q})}\right) + \delta\tau + \left(\frac{\beta}{\gamma}\right)\right]\cdot\exp\left\{-\gamma\left(\frac{p}{\pi(\boldsymbol{q})}\right)\right\}. \tag{16}$$

Substituting $y = px + y_z$, where $y_z \equiv \boldsymbol{q}'\boldsymbol{z}$ is the total expenditure on all goods other than x, into (14) and solving for p gives

$$p(x,y_z,\boldsymbol{q},\tau) = -\left(\frac{\alpha(\boldsymbol{q}) + \gamma y_z + \delta\tau - x}{\beta + \gamma x}\right)\cdot\pi(\boldsymbol{q}). \tag{17}$$

Then, substituting $p(x,y_z,q,\tau)$ into the quasi-indirect utility function gives the ***quasi-utility function*** (LaFrance; LaFrance and Hanemann) as

$$(18) \qquad \omega(x,y_z,q,\tau) = \left(\frac{\beta+\gamma x}{\gamma^2}\right)\cdot \exp\left\{\frac{\gamma[\alpha(q)+\gamma y_z+\delta\tau - x]}{\beta+\gamma x}\right\}.$$

I will use this expression for the quasi-utility function to demonstrate the precise cardinal scaling affect of the WCH for the linear demand model.

The compensated demand for x is found by differentiating (15) with respect to p,

$$(19) \qquad g(p,q,u,\tau) = \frac{\partial e(p,q,u,\tau)}{\partial p} = \left(\frac{\gamma\cdot\theta(q,u,\tau)}{\pi(q)}\right)\cdot \exp\left\{\gamma\left(\frac{p}{\pi(q)}\right)\right\} - \left(\frac{\beta}{\gamma}\right).$$

The right-hand-side vanishes if and only if

$$(20) \qquad \exp\left\{\gamma\left(\frac{p}{\pi(q)}\right)\right\} = \left(\frac{\beta\cdot\pi(q)}{\gamma^2\cdot\theta(q,u,\tau)}\right),$$

which implicitly defines $p^\circ(q,u,\tau)$. Now if we differentiate (15) with respect to τ, evaluate the derivative at $p^\circ(q,u,\tau)$, set the resulting expression equal to zero, and regroup some terms, we obtain the WCH partial differential equation as

$$(21) \qquad \frac{\partial\theta(q,u,\tau)/\partial\tau}{\theta(q,u,\tau)} = \frac{\gamma\delta}{\beta}.$$

The solution to (21) can be seen by inspection to be

$$(22) \qquad \theta(q,u,\tau) = \tilde{\theta}(q,u)\,e^{\gamma\delta\tau/\beta}.$$

Thus, the transformed expenditure function under the WCH is

$$(23) \qquad \tilde{e}(p,q,u,\tau) = \tilde{\theta}(q,u)\cdot \exp\left\{\gamma\left[\left(\frac{p}{\pi(q)}\right)+\left(\frac{\delta\tau}{\beta}\right)\right]\right\} - \left(\frac{\pi(q)}{\gamma}\right)\cdot\left[\alpha(q)+\beta\left(\frac{p}{\pi(q)}\right)+\delta\tau+\left(\frac{\beta}{\gamma}\right)\right],$$

the WCH transformed quasi-indirect utility function is

$$(24) \qquad \tilde{\varphi}(p,q,y,\tau) = \varphi(p,q,y,\tau)\cdot e^{-\gamma\delta\tau/\beta}$$

$$= \left(\frac{1}{\gamma}\right)\cdot\left[\alpha(q)+\beta\left(\frac{p}{\pi(q)}\right)+\gamma\left(\frac{y}{\pi(q)}\right)+\delta\tau+\left(\frac{\beta}{\gamma}\right)\right]\cdot \exp\left\{-\gamma\left[\left(\frac{p}{\pi(q)}\right)+\left(\frac{\delta\tau}{\beta}\right)\right]\right\},$$

and the WCH transformed quasi-utility function is

(25) $$\tilde{\omega}(x,y_z,q,\tau) = \omega(x,y_z,q,\tau)\,e^{-\gamma\delta\tau/\beta} = \left(\frac{\beta+\gamma x}{\gamma^2}\right)\cdot\exp\left\{\frac{\gamma[\alpha(q)+\gamma y_z-(\beta+\gamma\delta\tau)x/\beta]}{\beta+\gamma x}\right\}.$$

The first expression for $\tilde{\omega}(x,y_z,q,\tau)$ shows that the WCH produces a ***linear*** transformation of the preference map between the market good x and the Hicks composite commodity y_z, and the factor of proportionality is exponential in τ. Therefore, welfare measures for changes in τ are necessarily influenced in a nontrivial manner by the WCH. The second expression shows that the transformed direct preference map for x and y_z does indeed satisfy the WCH.[4]

Under the ANH for the linear demand model in (14), the structure of the expenditure function is

(26) $$\hat{e}(p,q,u,\tau) = \hat{\theta}(q,u)\cdot\exp\left\{\gamma\left(\frac{p}{\pi(q)}\right)\right\} - \left(\frac{\pi(q)}{\gamma}\right)\cdot\left[\alpha(q)+\beta\left(\frac{p}{\pi(q)}\right)+\delta\tau+\left(\frac{\beta}{\gamma}\right)\right],$$

while the quasi-indirect utility function has the form given in (16) above. Note that x and τ are globally Hicks neutral under the ANH, i.e., $\partial^2\hat{e}(p,q,u,\tau)/\partial p\partial\tau \equiv 0$. It is also easy to show that the good x is

---

[4] That this also implies that the overall utility function, $u(x,z,\tau)$, satisfies the WCH can be demonstrated as follows. Given the WCH, the indirect utility function satisfies

$$\tilde{v}(p,q,y,\tau) \equiv \tilde{\psi}(q,\tilde{\varphi}(p,q,y,\tau)).$$

Therefore, the relationship between the quasi-utility function and the utility function, can be written as (LaFrance and Hanemann)

$$\tilde{u}(x,z,\tau) \equiv \min\{\tilde{\psi}(q,\tilde{\omega}(x,q'z,\tau)) : q \ge 0\}.$$

Assuming $\tilde{u}(x,z,\tau)$ is differentiable in τ, it follows from the envelope theorem that

$$\frac{\partial\tilde{u}(x,z,\tau)}{\partial\tau} \equiv \frac{\partial\tilde{\psi}(q^*(x,z,\tau),\tilde{\omega}(x,q^*(x,z,\tau)'z,\tau))}{\partial\varphi}\cdot\frac{\partial\tilde{\omega}(x,q^*(x,z,\tau)'z,\tau)}{\partial\tau},$$

where $q^*(x,z,\tau)$ is the minimizing quantity-dependent price vector for all other goods. Because $\tilde{\psi}(q,\tilde{\varphi}(p,q,y,\tau))$ is monotonic in $\tilde{\varphi}$, it follows that $\partial\tilde{u}(x,z,\tau)/\partial\tau = 0$ if and only if $\partial\tilde{\omega}(x,q^*(x,z,\tau)'z,\tau)/\partial\tau = 0$.

weakly Hicks neutral from τ if and only if x is strictly Hicks neutral from τ for the linear demand model. As a result, this example conveniently illustrates the highly normative nature of both the weak complementarity and weak neutrality assumptions.

With some algebra the compensating variation measures $\tilde{c}$ and $\hat{c}$ for the change in τ from τ° to τ' obtained from (24) and (16), respectively, can be written as

$$\tilde{c} = \frac{\pi(q)}{\gamma^2}\cdot\left[\beta + \gamma x^\circ - (\beta + \gamma x')\cdot\left(1 - e^{\gamma\delta\Delta\tau/\beta}\right)\right], \tag{27}$$

$$\hat{c} = \frac{\pi(q)}{\gamma}\cdot\Delta x = \frac{\pi(q)}{\gamma}\cdot\delta\Delta\tau, \tag{28}$$

where $x^\circ = \alpha(q) + \beta\left(\frac{p}{\pi(q)}\right) + \gamma\left(\frac{y}{\pi(q)}\right) + \delta\tau^\circ$, $\Delta\tau = \tau' - \tau^\circ$, and $\Delta x = x' - x^\circ$.

The comparison between $\tilde{c}$ and $\hat{c}$ is greatly facilitated with the following definitions:

$$\text{(a) } \xi_p = \frac{\beta p}{\pi(q)x}; \text{ (b) } \xi_y = \frac{\gamma y}{\pi(q)x}; \text{ (c) } \xi_\tau = \frac{\delta\tau}{x}; \text{ (d) } w = \frac{px}{y}; \text{ (e) } \kappa = \frac{w\xi_y}{\xi_p}; \text{ and (f) } \eta = \xi_\tau\cdot\left(\frac{\Delta\tau}{\tau}\right).$$

The variables ξ_p, ξ_y, and ξ_τ are the demand elasticities for own price, income, and the nonmarket quality, respectively, and w is the budget share of the market good of interest. Given $\beta < 0$, i.e., a negatively-sloped ordinary demand, the variable κ is an index of concavity of the expenditure function with respect to p and must be greater than or equal to -1. To see this, write the own-price substitution term for x as

$$\frac{\partial x}{\partial p} + \frac{\partial x}{\partial y}\cdot x = \frac{\beta + \gamma x}{\pi(q)} = \frac{\beta}{\pi(q)}\cdot(1+\kappa). \tag{29}$$

This shows that $-1 \le \kappa$ and as κ approaches minus one from above, the expenditure function becomes less concave in p. Finally, the variable η measures the total percentage shift in the demand for the market good that is due to a percentage change in the nonmarket quality in the amount of $100\cdot(\Delta\tau/\tau)$.

The motivation for these definitions is that, with a little more algebra, it is possible to write the relative difference between $\tilde{c}$ and $\hat{c}$ as a simple function of the two variables κ and η,

$$\frac{\tilde{c} - \hat{c}}{\hat{c}} = \frac{(1+\kappa)\cdot\left(1 - e^{\kappa\eta}\right)}{\kappa\eta}. \tag{30}$$

It follows from this expression that the percentage difference between the compensating variation for the WCH and ANH hypotheses increases with: (a) an increase in the budget share for x; (b) an increase in the income elasticity of demand for x; (c) a decrease in the own-price elasticity of demand for x; and (d) an increase in the elasticity of demand with respect to the nonmarket quality and/or the percentage change in the nonmarket quality.

Figure 1 illustrates the situation where, without loss in generality, it is assumed that $\eta \geq 0$. In many typical applications to problems of valuing environmental improvements, the budget share (e.g., the proportion of income spent on fishing trips to a given lake) tends to be small, the income elasticity of demand is positive but modest, and except perhaps in cases of unique irreplaceable environmental assets, numerous close substitutes (other lakes and streams that can be fished) are available, so that the own-price elasticity of demand is relatively large. Under these conditions, in cases such as the demand for recreation visits, κ tends to be much closer to zero than minus one. But this is precisely the region where the percentage difference between the two welfare measures differ the most, ***for all values*** of the quality response term η.

This raises a serious question. How does one choose between $\tilde{c}$ and $\hat{c}$, or any of the uncountable number of alternative welfare measures, and defend that choice effectively? The rationale can not be based on convenience of the WCH - clearly the ANH is even more convenient since the specific WCH transformation of the preference map does not have to be derived. Nor can it be based on testable implications - none arise from either assumption or any of a very large class of competing alternatives. Assertions of low or zero non-use or existence values (McConnell; Mendelsohn 1984) are not compelling either, because simply put, such claims are merely assertions. Unfortunately, using the WCH (or any of its equally speculative cousins) is a bit like making up numbers and calling them welfare estimates.

4. CONCLUSIONS

This paper challenges a bit of conventional wisdom in applied welfare economics. The approach to valuing the benefits of environmental improvement or preservation of natural environments in which it is assumed that the welfare effects of changes in a nonmarket quality can be derived from observations on a set of purchased market goods has been shown to generate cardinal transformations of preferences which are not testable. The welfare measurements obtained from such practices therefore are arbitrary and can not be defended against an uncountable number of alternative assumptions that are each equally plausible and equally non-testable. At the very least, it is neither possible nor reasonable to assert that welfare measures calculated under any of these normative preference transformations contain any information about non-use or existence values. Pushing the matter a little further, the results of this paper suggest that this entire research strategy suffers from a serious weakness - ***there is <u>no behavioral hypothesis</u> underpinning weak complementarity or any of its alternatives, yet these hypotheses have significant normative impacts on the estimated values of environmental improvements.***

REFERENCES

Arrow, K. and A. Fisher. "Environmental Preservation, Uncertainty, and Irreversibility." *Quarterly Journal of Economics* 88 (1974): 312-319.

Bockstael, N., W. M. Hanemann, and C. Kling. "Estimating the Value of Water Quality Improvements in a Recreational Demand Framework." *Water Resources Research* 23 (1987): 951-960.

Bockstael, N. and C. Kling. "Valuing Environmental Quality When Quality is a Weak Complement to a Set of Goods." *American Journal of Agricultural Economics* 70 (1988): 654-662.

Bockstael, N. and K. McConnell. "Welfare Measurement in the Household Production Framework." *American Economic Review* 73 (1983): 806-814.

Braden, J. and C. Kolstad, ed. *Measuring the Demand for Environmental Quality* Amsterdam: North-Holland, 1991.

Cicchetti, C. and A. M. Freeman III. "Option Demand and Consumer Surplus: Further Comment." *Quarterly Journal of Economics* 85 (1971): 528-539.

Diewert, W. E. "Hicks' Aggregation Theorem and the Existence of a Real Value-Added Function." *Production Economics: A Dual Approach to Theory and Applications, Vol. 2*, ed. M. Fuss and D. McFadden. New York: North-Holland, 1978: 17-51.

Epstein, L. "A Disaggregate Analysis of Consumer Choice Under Uncertainty." *Econometrica* 47 (1975): 877-892.

Epstein. L. "Integrability of Incomplete Systems of Demand Functions." *Review of Economic Studies* 49 (1982): 411-25.

Freeman, A. M. III. *The Benefits of Environmental Improvement* Baltimore: Johns Hopkins University Press for Resources for the Future, 1979.

Hausman, J. "Exact Consumer's Surplus and Deadweight Loss." *American Economic Review* 71 (1981): 662-672.

LaFrance, J. "Linear Demand Functions in Theory and Practice." *Journal of Economic Theory* 37 (1985): 147-166.

LaFrance, J. and W. M. Hanemann. "The Dual Structure of Incomplete Demand Systems." *American Journal of Agricultural Economics* 71 (1989): 262-274.

Larson, D. "Recovering Weakly Complementary Preferences." *Journal of Environmental Economics and Management* 21 (1991): 97-108.

Larson, D. and P. Flacco. "Measuring Option Prices from Market Behavior." *Journal of Environmental Economics and Management* 22 (1992): 178-198.

Mäller, K.-G. *Environmental Economics: A Theoretical Inquiry* Baltimore: Johns Hopkins University Press for Resources for the Future, 1974.

Mendelsohn, R. "Modeling the Demand for Outdoor Recreation." *Water Resources Research* 23 (1987): 961-967.

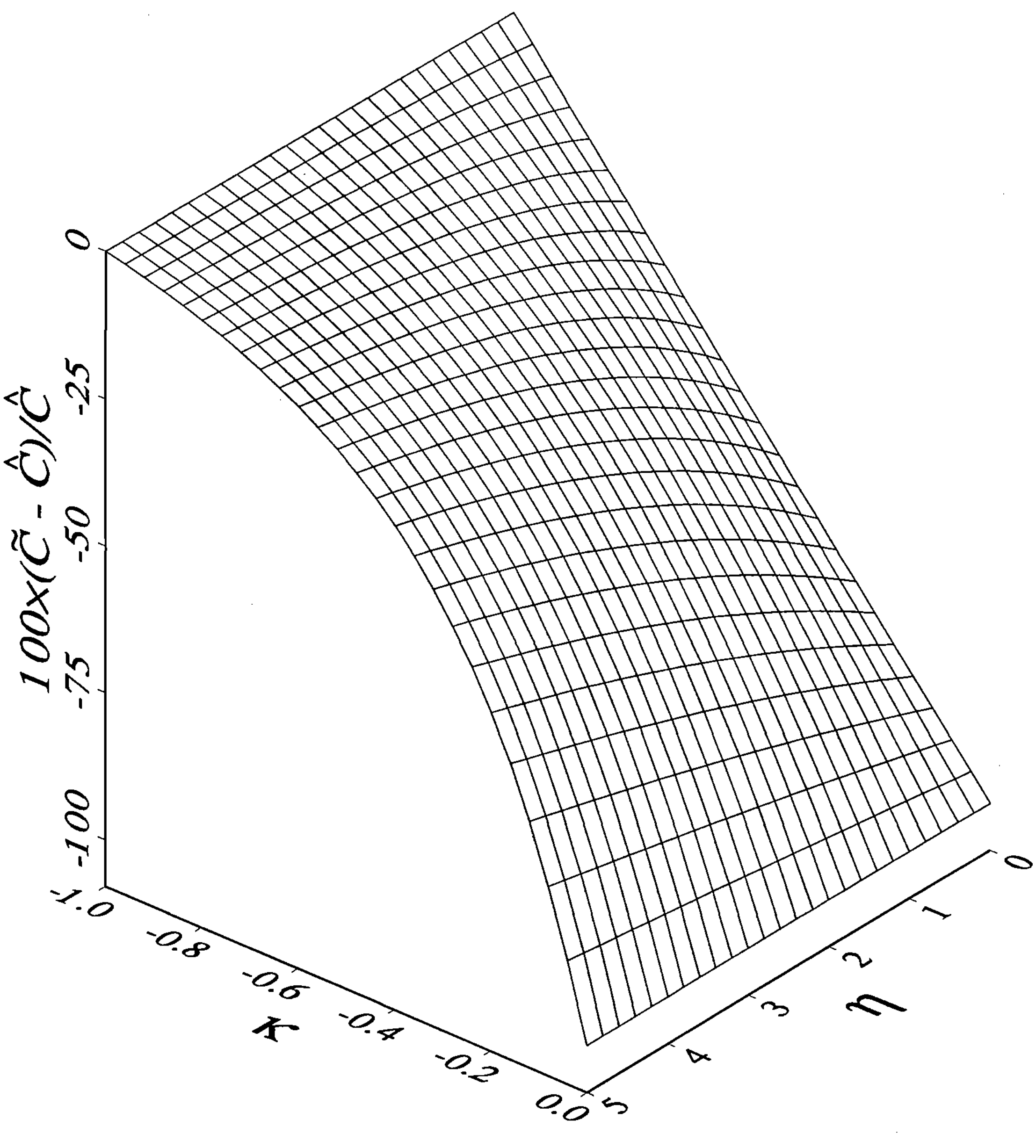
Smith, V. K. and W. Desvousges. "The Generalized Travel Cost Model and Water Quality Benefits: A Re-consideration." *Southern Economic Journal* 52 (1985): 371-381.

Smith, V. K., W. Desvousges, and M. McGivney. "Estimating Water Quality Benefits: An Econometric Analysis." *Land Economics* 59 (1983): 259-278.

Willig, R. "Incremental Consumer's Surplus and Hedonic Price Adjustment." *Journal of Economic Theory* 17 (1978): 227-253.

Figure 1. Weak Complementarity vs. Weak Neutrality.

Percentage Difference in Compensating Variation



DOUBLE HURDLE COUNT DATA MODELS FOR TRAVEL COST ANALYSIS

by

J.S. Shonkwiler
Professor
Department of Agricultural Economics
University of Nevada
Reno, Nevada 89557
(702) 784-1341

DOUBLE HURDLE COUNT DATA MODELS FOR TRAVEL COST ANALYSIS

Recreation demand modeling is an important element of natural resource planning. Behavioral responses and valuations of recreationists are often used as components of benefit-cost analysis or environmental impact assessment. The travel cost model which defines a demand function for recreation sites has been employed by economists since the early 1960's (Smith, 1989). Yet a number of theoretical and empirical problems encompass the travel cost model. These include issues involving the count data structure of the dependent demand quantity, assumptions regarding the structure of the demand decision relative to corner solutions and hurdles to consumption, and the inclusion of substitute sites in the empirical specification.

Count data travel cost models have become increasingly more common (Creel and Loomis, 1990; Hellerstein, 1991; Englin and Shonkwiler, 1993) as economists have recognized that travel cost studies based on participant information are subject to the fact that each respondent will report a discrete number of trips. Because trip demand cannot be negative, the data set is censored at zero and failure to account for censoring leads to biased estimation. The application of count data estimators to the travel cost model thus is a logical extension to accommodate the particular properties of trip data. In view of the recent work of Hellerstein and Mendelsohn, who provide theoretical foundations for linking the empirical count estimator to the individual consumer's underlying optimization problem, it is clear how to perform welfare analysis on the basis of a count demand model for trips.

There does appear to be some confusion regarding the interpretation of welfare measures when specialized forms of count data estimators are used. Truncated count data estimators have been considered by Creel and Loomis and Yen and Adamowicz in the analysis of recreation demand. In both studies, the consumer surplus measures derived are likely misinterpreted due to the fact that the underlying probability mass function was not established. It is important that consumer surplus be properly derived in the truncated count model, because more complicated models of recreationist decision making can be based on it. Specific interest is focused on hurdle types of models, as evidently first suggested by Mullahy in the count data context.

With regard to corner solutions and hurdles to consumption, Pudney has pointed out that individuals may make non-marginal changes by switching from one behavioral regime to another. The two types of responses involved are at the intensive margin where consumers of the good are motivated to consume marginally more and at the extensive margin where people may either enter or leave the market entirely. Individuals with zero consumption may be at a corner solution such that a price reduction (income increase) may lead to non-zero levels of consumption. Alternatively, zero consumption may represent behavior which is robust to changing economic variables, reflecting instead a choice set influenced by endowments or physical capabilities.

While econometric models developed by Mullahy may provide insight into the treatment of zero consumption as the consequence of a corner solution to the conventional utility maximization problem, they are unsatisfactory in terms of their ability to distinguish situations where desired consumption may be positive but observed consumption is recorded as zero. Using a specification developed by Blundell and Meghir in the continuous consumption case, this paper will extend their results to the discrete consumption case. A multivariate count data model is also developed to permit dependence between stages of decision making.

The paper proceeds by reviewing calculation of consumer surplus for the standard count data model. Welfare measures for the truncated count data model are derived. These measures are related to the single hurdle model (Mullahy) of consumption. Next, the double hurdle model of consumption behavior is specified both under independence and dependence. Again, appropriate welfare measures are derived. Finally, these models are illustrated using a stylized model of angler demand for a Nevada lake.

Consumer Surplus in the Count Data Model

Begin by assuming that the i^{th} potential user of a specific recreation site has been randomly drawn from a relevant population. Let y_i denote the number of visits to the recreation site made by the individual. If

(1) $$pr(Y_i = y_i) = e^{-\lambda_i} \lambda_i^{y_i} / y_i!, \qquad y_i = 0,1,2,3,...$$

where Y_i is a potential integer outcome, then it is well known that $E(Y_i) = \lambda_i$. When λ_i is parameterized as

(2) $\lambda_i = \exp(\alpha_i + \beta p_i),\qquad \beta < 0$

so that the price parameter is made explicit, consumer surplus (or more formally the expected value of consumer surplus) is calculated as

(3) $\int_{p_i}^{\infty} \lambda_i dp = -\lambda_i / \beta$

In empirical applications, estimated values of the parameters are used to calculate (3) for each individual and an average consumer surplus measure is calculated.

If the random sample consists only of users, then $q_i = y_i | y_i > 0$. The use of q_i denotes that the underlying population of potential users from which y_i was drawn has been truncated. It is important to recognize that the sample q of which q_i is an element can only be used to provide information about the population of users, it can tell us nothing about the population of users and non-users unless the probability mass function in (1) has been empirically verified. Since q is a subset of y (the sample of the population of potential users), knowledge of y can always be used to make inferences concerning users, however, knowledge of q cannot be used to make inferences about users and non-users because there is no means to establish (1) as the underlying probability mass function.

The probability function for the j[th] individual in the sample of users given the Poisson parameter ψ is

(4) $pr(Q_j = q_j) = \dfrac{e^{-\psi_j} \psi_j^{q_j} / q_j!}{1 - e^{-\psi_j}}\qquad q_j = 1,2,3.....$

and $E(Q_j) = \psi_j / (1 - e^{-\psi_j})$, i.e. the positive Poisson distribution. Because the denominator in (4) can be interpreted as the probability that q_i is greater than zero, analysts have been led to interpret $e^{-\psi_i}$ as the probability of non-use. The discussion of Grogger and Carson (p. 230) regarding this issue fails to stress their assumption that they know with certainty the underlying distribution is Poisson with the identical location parameter $(\lambda_j = \psi_j, \forall_j)$. In general, the sample q can tell us nothing about non-users because it contains no information about non-users. This probability, $e^{-\psi_i}$, has no economic interpretation relating to decisions to enter or leave the market. Instead, the denominator in (4) accounts for the correction necessary when Q_i is prohibited from being zero so that (4) is a valid probability mass function.

To calculate expected consumer surplus when $\psi_j = \exp(\alpha_j + \beta p_j)$ requires the determination of a choke price which drives quantity to zero. Note that[1]

(5) $$\lim_{p \to \infty} E(Q|Q>0) = \lim_{p \to \infty} \frac{\exp(\alpha+\beta p)}{1-e^{-\exp(\alpha+\beta p)}} = 1.$$

That is, demand cannot be driven to zero for the truncated model since quantities are constrained to be non-zero as a consequence of the probability model. Thus, we can integrate under the truncated demand curve up to the point where $Q = 1$. This is shown in Figure 1. For the truncated model, the consumer surplus associated with $Q-1$ trips is given by the improper integral

(6) $$\lim_{c \to \infty} \int_{P_0}^{c} (Q-1)dp = \lim_{c \to \infty} \left[\beta^{-1} \ln\left(e^{\exp(\alpha+\beta p)} - 1\right) - P\right]_{P_0}^{c}$$

Because the limit exists[2] the surplus measure is

(7) $$CS(Q-1) = \frac{\alpha}{\beta} - \beta^{-1} \ln\left(e^{\exp(\alpha+\beta P_0)} - 1\right) + P_0.$$

By convention, the discrete nature of the count form of quantity demanded is ignored when calculating consumer surplus. To develop the adjustment to equation (7) required to derive a continuous consumer surplus measure for users, an analogy to how the consumer surplus for a user in the untruncated model is presented. Refer to Figure 2a. where the price and quantity axis have been transposed and notice that if $\lambda = \exp(\alpha + \beta P)$ then P^*, the price at which Q=1, is given by $-\alpha/\beta$. This makes area $C = -\alpha/\beta - P_0$. The area of $B + D = -\lambda/\beta + \alpha/\beta + P_0$. Now consider figure 2b. where truncation makes the area of $B' + D'$ the quantity given in equation (7). Because $\lim_{\psi \to \infty} -\beta^{-1} \ln\left(e^{\psi} - 1\right) = -\psi/\beta$ we see

---

[1]Let $y = e^{-\exp(\alpha+\beta p)}$ then the limit in (5) can be expressed as $\lim_{y \to 1} -\frac{\ln y}{1-y} = \lim_{y \to 1} -\frac{y^{-1}}{-1} = 1$ when l'Hopital's rule is applied.

[2]To show that $\lim_{p \to \infty} \beta^{-1} \ln\left(e^{\exp(\alpha+\beta p)} - 1\right) - p$ exists, let $y = e^{\alpha+\beta p}$ and consider $\lim_{y \to 0} \beta^{-1}\left(\ln\left(e^{y} - 1\right) + \alpha - \ln y\right) = \alpha/\beta$ given that $\lim_{y \to 0} \ln\left(e^{y} - 1\right) - \ln y$ can be written $\lim_{y \to 0} \ln\left(y + \frac{y^2}{2!} + \frac{y^3}{3!} + \ldots\right) - \ln y = 0$ since the infinite series converges.

that the areas of $B+D$ and $B'+D'$ have identical forms for large location parameters. This suggests adjusting equation (7) by adding in the area C which gives

(7') $$CS(\text{user}) = -\beta^{-1}\ln\left(e^{\gamma} - 1\right)$$

Equation (7') now permits calculation of the consumer surplus of users and consequent generalization to the population of <u>users</u>. To make statements regarding the consumer surplus of <u>potential users</u> requires an additional piece of information or an assumption regarding the percentage of potential users who are actual users. Then the consumer surplus of a potential user may be calculated as the consumer surplus of a corresponding user multiplied by the probability of use.

Hurdle Count Data Models

It is known that non-consumption of a particular site is a commonly observed phenomenon having both economic and statistical implications. Henceforth assume that y represents a sample of all potential users. Let y_i denote the number of visits the i$^{\text{th}}$ (potential) user of a specific recreation site has made over a season. Define two vectors of variables that influence the individual's decisions as x_i and z_i, with the former containing mainly economic variables (prices, income) and the latter containing mainly demographic variables (age, gender, marital status, etc.). This decomposition of explanatory variables is in the spirit of Pudney who suggests separating economic variables from personal characteristics which shape tastes. Let D_i represent the latent decision to consume such that consumption is zero if $D_i \leq 0$. Specify

(8) $$E(D_i) = \theta_i = \exp\left(z_i'\gamma\right)$$

where γ is an unknown vector of parameters.

If consumption is positive, then it is assumed that observed consumption equals desired consumption

(9) $$y_i = y_i^{*} \quad \text{with}$$

$$E\left(y_i^{*}\right) = \lambda_i = \exp\left(\alpha_i + \beta p_i\right)$$

The sequential (single hurdle) model then has a probability density function of the form

$$\text{(10)}\qquad \begin{array}{l} \Pr(D \le 0),\ \text{if } y = 0 \\ \Pr(y|y > 0)\Pr(D > 0),\ \text{if } y > 0 \end{array}$$

where the observational subscripts have been suppressed. The likelihood function in the case of Poisson specifications is (Mullahy)

$$\text{(11)}\qquad \prod_{y=0} \exp(-\theta) \prod_{y>0} (1-\exp(-\theta))\lambda^{y} / \left[(\exp(\lambda)-1)y!\right]$$

For this model

$$\text{(12)}\qquad E(y|y>0) = \frac{\lambda}{1-e^{-\lambda}} \text{ and } E(y) = \frac{\lambda\left(1-e^{-\theta}\right)}{1-e^{-\lambda}}.$$

Assuming that θ does not depend on p, then consumer surplus for the population is proportional to consumer surplus of the users and the formula in (7') will be required.

Consequently, the Mullahy model requires rather complicated calculations to determine consumer surplus for the population. Also consider that while this model may ameliorate the over-dispersion commonly observed in Poisson count data models, it is unattractive because the decision to consume is independent of the level of consumption. This condition stems from the fact that the information matrix is block diagonal in terms of the two regimes and equivalent maximum likelihood estimates may be obtained by separately estimating a binary choice model and a truncated Poisson.

The double hurdle model (without dependence) specifies the probability of a zero observation as

$$\text{(13)}\qquad \Pr(y^{*} \le 0) + \Pr(y^{*} > 0)\Pr(D \le 0).$$

No consumption will be observed if desired consumption is non-positive, or, if desired consumption is positive, an additional hurdle $(D \le 0)$ may prevent consumption. The probability of a positive observation is

$$\text{(14)}\qquad \Pr(y^{*} > 0)\Pr(y^{*}|y^{*} > 0)\Pr(D > 0)$$

so that the Poisson likelihood becomes

$$(15)\qquad \prod_{y=0}\left[\exp(-\lambda)+\left(1-\exp(-\lambda)\right)\exp(-\theta)\right]\cdot \prod_{y>0}\left(1-\exp(-\theta)\right)\exp(-\lambda)\lambda^{y}/y!$$

under the assumption that $y = y^{*}$ if $y^{*} > 0$ and $D > 0$. This model incorporates two mechanisms for generating zeros, one intended to represent a fundamental non-economic decision and the other representing an ordinary corner solution (Pudney).

In the case of the double hurdle model, we have[3]

$$(16)\qquad E(y|y>0)=\frac{\lambda}{1-e^{-\lambda}} \text{ and } E(y)=\lambda\left(1-e^{-\theta}\right)$$

So an attractive feature of this model is the ease by which consumers surplus can be calculated for the population when θ does not depend on p. In fact, the average consumer surplus per trip for the population is simply estimated as $-\beta^{-1}\left(1-e^{-\theta}\right)$.

Next, consider the relationship between the two hurdles. If those two mechanisms are correlated, a bivariate double hurdle model results. The two regimes may be represented for count data as

$$(17)\qquad \Pr(y=0)=\Pr(y^{*}=0)+\sum_{i=1}^{\infty}\Pr(y^{*}=i,D=0)$$

and

$$\Pr(y>0)=\sum_{j=1}^{\infty}\Pr(y^{*}=y,D=j)$$

Using Holgate's bivariate Poisson model (Johnson and Kotz), the likelihood for the double hurdle count model under dependence may be expressed as[4]

---

[3]The second relation follows since $\Pr(y>0)=\left(1-e^{-\lambda}\right)\left(1-e^{-\theta}\right)$.

[4]The marginal distributions of y^* and D are Poisson with parameters $\lambda+\xi$ and $\theta+\xi$, respectively. To derive $\Pr(y=0)$ note that

$$\begin{aligned}\sum_{i=1}^{\infty}\Pr(y^*=i,D=0)&=\sum_{i=0}^{\infty}\Pr(y^*=i,D=0)-\Pr(y^*=0,D=0)\\&=\Pr(D=0)-\Pr(y^*=0,D=0).\end{aligned}$$

$$\prod_{y=0}\left[\exp(-\lambda-\xi)+\exp(-\theta-\xi)-\exp(-\lambda-\theta-\xi)\right]\cdot$$

(18)
$$\prod_{y>0}\left[\exp(-\lambda-\xi)(\lambda+\xi)^{y}/y!-\exp(-\lambda-\theta-\xi)\lambda^{y}/y!\right]$$

where ξ represents the covariance between the two Poisson processes. Note that (18) collapses to (15) when ξ is zero, hence independence is nested relative to dependence.

It can be shown that[5]

(19) $$E(y|y>0)=\frac{\lambda(1-\Pr(D=0))+\xi}{1-\Pr(y=0)}=\frac{\lambda+\xi-\lambda e^{-\theta-\xi}}{1-e^{-\lambda-\xi}-e^{-\theta-\xi}+e^{-\lambda-\theta-\xi}}$$

so that

(20) $$E(y)=\lambda+\xi-\lambda e^{-\theta-\xi}$$

From (20) it can be seen that consumer surplus is readily calculated as long as ξ and θ do not depend on p. However, it may be reasonable to parameterize the covariance according to

(21) $$\xi_{t}=\exp(\phi_{t}+\delta p_{t}),\qquad \delta<0$$

In this case we have

(22) $$CS=-\gamma/\beta-\xi/\delta-e^{-\theta}\int_{P_0}^{\infty}\lambda e^{-\xi}dp$$

and by a change of variables the integral can be expressed in terms of an incomplete gamma function. By using the infinite series representation of the incomplete gamma function, we see

(23) $$CS=-\lambda/\beta-\xi/\delta+\delta^{-1}\exp(\alpha-\theta-\beta\phi\delta^{-1})\sum_{j=0}^{\infty}\frac{(-1)^{j}\xi^{j+\beta/\delta}}{j!(j+\beta/\delta)}$$

---

[5]In a similar vein the $\Pr(y>0)$ can be expressed

$$\sum_{j=1}^{\infty}\Pr(y^*=y,D=j)=\sum_{j=0}^{\infty}\Pr(y^*=y,D=j)-\Pr(y^*=y,D=0)$$
$$=\Pr(y^*=y)-\Pr(y^*=y,D=0).$$

$$E(y|y>0)=\sum_{i=1}^{\infty}\frac{i\Pr(y=i)}{1-\Pr(y=0)}$$
$$=\frac{\sum_{i=1}^{\infty}i\{\Pr(y^*=i)-\Pr(y^*=i,D=0)\}}{1-\Pr(y^*=0)-\Pr(D=0)+\Pr(y^*=0,D=0)}$$

The numerator then may be written as $E(y^*)-E(y^*|D=0)\Pr(D=0)$.

Empirical Analysis

The data to be analyzed were collected in 1988 through a cooperative effort by the Nevada Department of Wildlife and the Department of Agricultural Economics, University of Nevada. Residents holding valid Nevada fishing licenses were surveyed. Respondents were asked to supply information to questions regarding household characteristics (number of family members, age and gender of family members, education levels), household income, and expenditures on equipment, guide and taxidermy services. Respondents were also asked to provide single and multiple site trip information regarding visits to a number of Nevada lakes and reservoirs.

For the purposes of this study, visits to Pyramid Lake by northern Nevada anglers are analyzed in order to infer its use value to area anglers. Pyramid Lake is located 35 miles northeast of Reno and is highly regarded for its fly fishing. The sample used consists of 659 anglers of which 67 (10.17%) report at least one, but no more than twelve, visits to Pyramid Lake during the 1988 season. These 67 users reported an average 4.84 trips whereas the average number of trips by all anglers in the sample is only .47. Variables which comprise the z matrix are an intercept, a dummy variable denoting the respondent's age as over 50 (Age), and a dummy variable denoting ownership of a boat (Boat). The x matrix contains an intercept, a price variable calculated as respondent's round trip mileage from the lake multiplied by \$.25/mile, and the respondent's household income in 1,000's of dollars.

Estimation results for the three hurdle models are presented in Table 1. The reader is cautioned to regard estimated parameters and corresponding standard errors with some suspicion given the sensitivity of the truncated Poisson to overdispersion (Grogger and Carson). However, Gurmu has pointed out that the positive Poisson density is a member of the linear exponential family so according to the results of Gourieroux et al., it should provide consistent estimators of the conditional mean irrespective of distributional misspecification. Furthermore, the estimated standard errors of the parameters have been corrected using Theorem 3 of Gourieroux et al., therefore, comparisons between models may be informative.

The single hurdle model treats the λ regime as an independent truncated regression on users, thus the parameters on x for the stand-alone truncated model are also known. The double hurdle model allows economic variables to affect both the non-consumption and consumption decisions. As a result, the price coefficient is almost 50 percent larger (absolutely) than that of the single hurdle model. The double hurdle with dependence model suggests that ξ is significantly greater than zero with a likelihood ratio test of 14.6 which is distributed as χ^2 with three degrees of freedom. Covariance between the decision making processes increases with proximity to the lake and ownership of a boat.

Average consumer surplus measures are presented in Table 2 for the three models. Apparently, users value the lake at about $250 per season. Interestingly, the double hurdle models give similar results for all potential users who are estimated to value the lake at about $30 per season. Given a population of 40,000 anglers in northern Nevada, Pyramid Lake represents a use value of about $1,200,000 per season to area anglers.

Conclusions

Economists have used the travel cost model of recreation demand for several decades. The fact that trip demand is measured as a discrete, non-negative variable has led to the recent use of count data estimators. The theoretical foundation for the use of count data models in welfare analysis has been developed for analyzing individual recreation data by Hellerstein and Mendelsohn, who advocate their application to measuring the non-market value of recreation sites. Non-consumption of a particular site is a commonly observed phenomenon having both economic and statistical implications. Studies which analyze only users are limited in that a truncated sample cannot be used to estimate benefits for potential users.

Unless participation behavior is carefully described and captured empirically, it is likely that subsequent valuations may be distorted. Since fishing requires specialized equipment and physical activity, certain lakes or reservoirs may not be visited due to lack of a boat or inability to reach a specific locations. Investigators need to recognize the importance of alternative representations of the decision making process irrespective of the fact that consumer surplus measures have non-standard forms. This study has provided a blueprint for such calculations. Extensions to the truncated negative binomial model follow analogously, although it is unclear whether a closed form expression for the surplus measure exists.

References

Blundell, R. and C. Meghir. "Bivariate Alternatives to the Tobit Model." Journal of Econometrics, 34(1987): 179-200.

Creel, M. and J. Loomis. "Theoretical and Empirical Advantages of Truncated Count Data Estimators." American Journal of Agricultural Economics, 72(1990): 434-441.

Englin, J. and J.S. Shonkwiler. "Estimating Social Welfare Using Count Data Models." Review of Economics and Statistics, (forthcoming).

Gourieroux, C., A. Monfort and A. Trognon. "Pseudo Maximum Likelihood Methods: Theory." Econometrica, 52(1984): 681-700.

Grogger, J.T. and R.T. Carson. "Models for Truncated Counts." Journal of Applied Econometrics, 6(1991): 225-238.

Gurmu, S. "Tests for Detecting Overdispersion in the Positive Poisson Regression Model." Journal of Business and Economic Statistics, 9(1991): 215-222.

Hellerstein, D.M. "Using Count Data Models in Travel Cost Analysis with Aggregate Data." American Journal of Agricultural Economics, 73(1991): 860-866.

Hellerstein, D.M. and R. Mendelsohn. "A Theoretical Foundation for Count Data Models." American Journal of Agricultural Economics, 75(1993): 604-611.

Holgate, P. "Estimation for the Bivariate Poisson Distribution." Biometrika, 51(1964): 241-245.

Johnson, N. and S. Kotz. Discrete Distributions. New York: John Wiley, 1969.

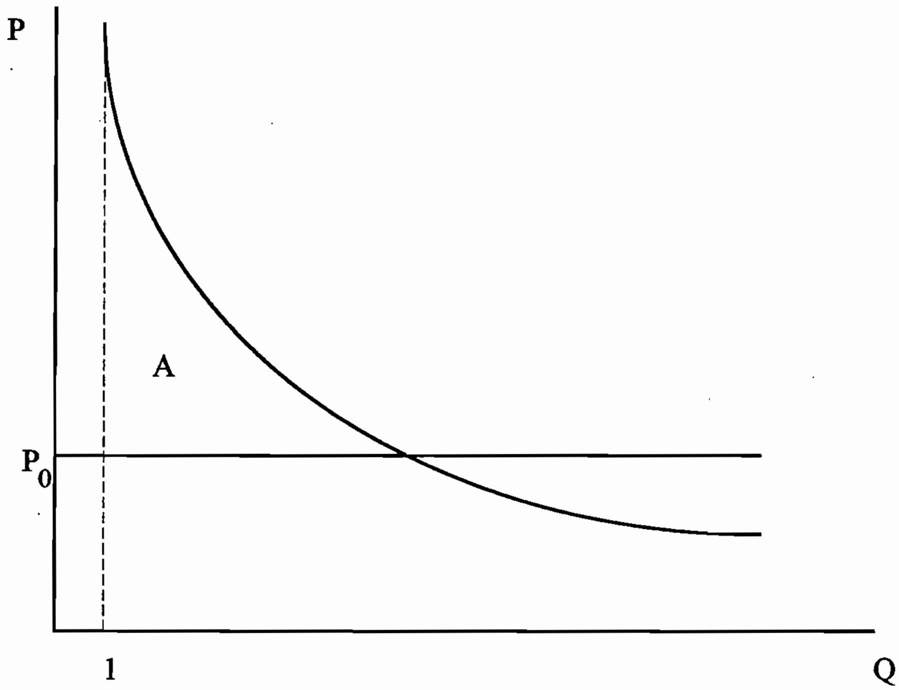
Mullahy, J. "Specification and Testing of Some Modified Count Data Models." Journal of Econometrics, 33(1986): 341-365.

Pudney, S. Modeling Individual Choice. New York: Basil Blackwell, 1989.

Smith, V.K. "Taking Stock of Progress with Travel Cost Recreation Demand Models: Theory and Implementation." Marine Resource Economics, 6(1989): 279-310.

Yen, S.T. and W.L. Adamowicz. "Statistical Properties of Welfare Measures from Count-Data Models of Recreation Demand." Review of Agricultural Economics, 15(1993): 203-216.

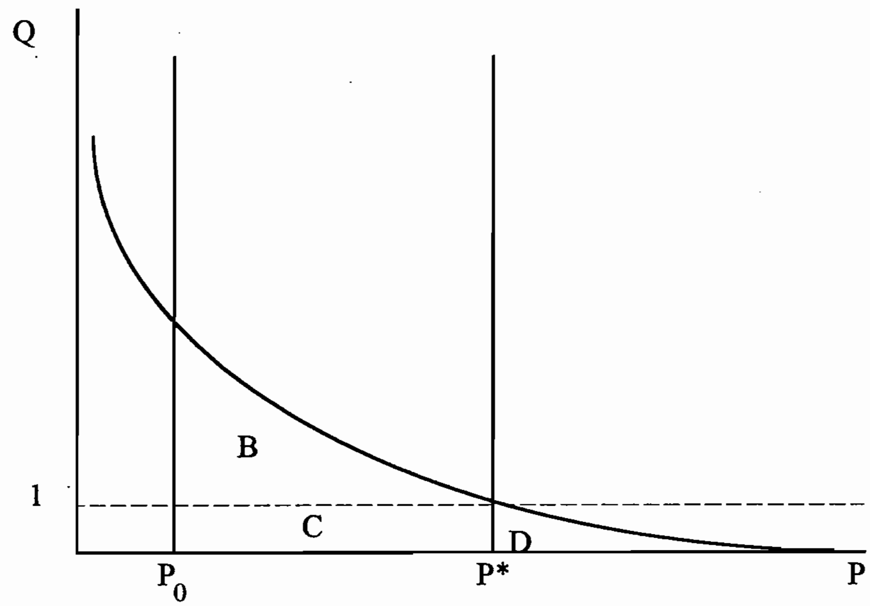
Figure 1. Consumer's Surplus Associated with Q-1 Trips.



$$A = \int_{P_0}^{\infty} (Q-1)dp$$

Figure 2. Comparison of Untruncated and Truncated Demands.

a.



b.

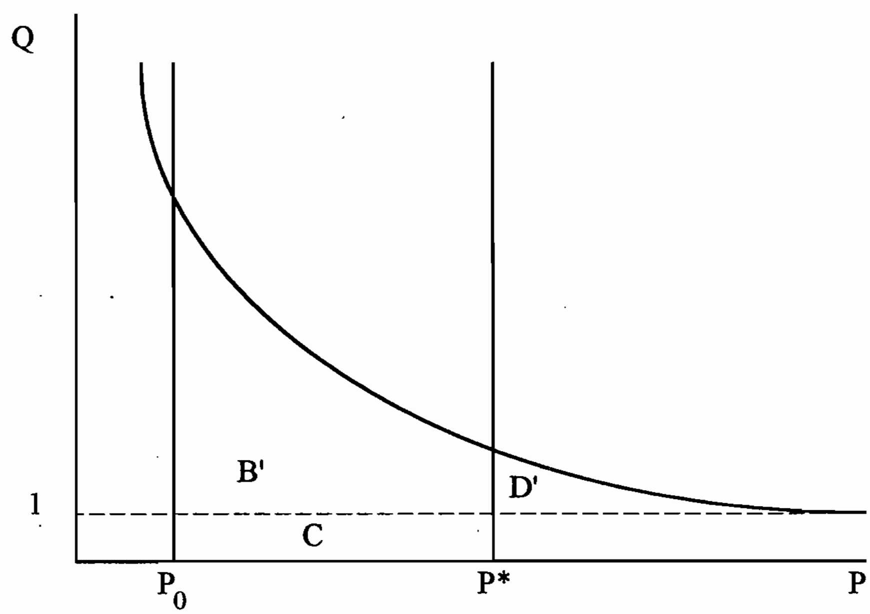


Table 1. Estimation Results.

		Model					
		Single Hurdle		Double Hurdle		Double Hurdle with Dependence	
Regime	Variable	Parameter	Std. Error	Parameter	Std. Error	Parameter	Std. Error
θ	Intercept	-2.520	.215	-2.453	.225	-2.782	.371
θ	Age	-.839	.265	-.836	.265	-1.503	.603
θ	Boat	1.142	.261	1.135	.261	.935	.546
λ	Intercept	1.960	.316	2.188	.422	1.842	.424
λ	Price	-.015	.008	-.022	.012	-.016	.011
λ	Income	.003	.003	.003	.003	.003	.004
ξ	Intercept					-.894	.913
ξ	Price					-.080	.017
ξ	Boat					1.557	.713
log likelihood		-379.4		-377.5		-370.2	

Table 2. Consumer Surplus Estimates.

	Model		
Group	Single Hurdle	Double Hurdle	Double Hurdle with Dependence
Users	313.82	210.30	240.22
All	44.55	31.47	29.56

A COMPARISON OF METHODS FOR ESTIMATING RECREATION DEMAND MODELS WITH COLLINEAR FACILITIES

by

Brian A. Roach[1]

and

Frank A. Ward[2]

ABSTRACT

Estimating natural resource values with travel cost demand models is limited by high collinearity when facilities are engineered in constant proportions across sites. Using Monte Carlo simulation methods with actual site facility data, the performance of a facility index is compared against three standard econometric approaches for dealing with collinearity.

[1] University of California, Davis

[2] New Mexico State University

A COMPARISON OF METHODS FOR ESTIMATING RECREATION DEMAND MODELS WITH COLLINEAR FACILITIES

I. INTRODUCTION

This paper is based on work being conducted under contract with the U. S. Army Corps of Engineers that estimates travel cost models (TCM) for Corps reservoirs. The benefits of changes in each of several classes of site facility variables are of considerable interest to Corps planners. However, high multicollinearity among facility variables results when several variables are included in a TCM regression. This collinearity problem results from the fact that as reservoirs increase in size they are typically built with more facilities in approximately constant proportions. The result of this engineered collinearity is unstable parameter estimates in the demand and benefits models for changes in the level of any one facility.

Recent simulation work in the recreation demand literature has explored the implications of various functional forms (Kling, 1987). This article seeks to further exploit the potential of simulations to provide additional insight into model construction and interpretation.

Results presented below report on a Monte Carlo simulation experiment which compares the effectiveness of different methods of estimating parameters for each facility variable in the presence of high multicollinearity. Several widespread remedies for dealing with collinearity are compared with a method developed for this study in which a facility index is constructed in a two-step process. The findings may have implications for model construction when several highly correlated policy variables must be included in a regression equation.

II. BACKGROUND

The simulation experiments presented report on four approaches for dealing with high multicollinearity. The first approach is to simply use OLS with the full set of regressors. Since OLS tends to produce unstable coefficients in the presence of high collinearity, OLS results with all the independent variables may be unacceptable for some policy purposes. This is especially the case when collinearity leads to coefficients with theoretically wrong signs.

The second approach follows one common method for dealing with high collinearity, i.e. dropping some independent variables. This results in the well-known omitted variable problem if theory supports the inclusion of the dropped variables. Also, dropping some variables restricts the potential to conduct policy analysis on a full set of variables controllable by policy.

The third approach included in the experiment is principal components analysis (PCA). PCA produces biased estimators which may have lower mean squared errors (MSE) than OLS coefficients. MSE is defined as the sum of an estimate's variance plus its bias squared (Mendenhall, et. al., 1990, p. 339). With high collinearity, a large proportion of the total variance among the independent variables may be explained by use of PCA by using only a few linear combinations of the variables. PCA has been criticized because it is not based upon a relationship between the independent and dependent variables. Wetzstein and Green (1978) used PCA in an analysis of wilderness area visitation in California.

The fourth approach developed for this paper for application to recreation benefits analysis is a facility index. The general specification of the index assumes a model:

$$Y = f(Z_1, \ldots, Z_m, X_1, \ldots, X_n, \epsilon) \tag{1}$$

where the X site facility variables are highly correlated with each other. The Z independent variables are other predictors of recreation demand, such as price, income and various demographic variables. These variables are typically much less correlated with each other and with the site facilities.

If one adopts the method of estimation by fitting the following model:

$$Y = f(Z_1, \ldots, Z_m, X_1, \epsilon') \tag{2}$$

an omitted variable problem clearly exists if the other X variables are theoretically correct. The first-stage of the index method developed for this paper involves estimating n regressions, each with all Z variables, but with and a different X variable, to obtain an (n*1) vector of biased coefficients on the X variables.

The rationale behind the use of our index is that the (n*1) vector of coefficients can be deflated or inflated closer to their true values through a second-stage regression. The vector of (n*1) biased coefficients are used as weights to construct a new independent variable which is a combination of the X variables. In a linear model, the index would be a linear combination using the first-stage biased coefficients as weights. The second stage model includes all Z variables as well as the index. The estimated coefficient on the index would then be combined with the first-stage biased coefficients to obtain a final estimate for the coefficient on each X variable.

The index proposed in this paper is applicable only when the magnitude of the bias on all n coefficients is similar. Thus, if the negative coefficients are biased downwards and the positive coefficients are biased upwards, then the index may be useful. This is because each first-stage coefficient is affected equally by the estimated second-stage coefficient on the index. We explore the advantages and disadvantages of such an index using a data set beset with high multicollinearity.

III. DATA SOURCES AND MODEL SPECIFICATION

The data for this paper were collected from records on 10 Corps of Engineer reservoirs in the Sacramento California District. Information on substitute water-based recreation sites was collected from all sites with more than 500 water surface acres of public recreation sites in California, Oregon, and Nevada. A zonal TCM was developed by aggregating to the county level with 348 total observations. The dependent variable, $VISITS_{ij}$ is the number of annual visits from county i to recreation site j. The independent variables included in the analysis and simulation experiments are:

POP_i	The population of county i.
$TOTCOST_{ij}$	The sum of travel costs and time opportunity costs from county i to site j. Time was valued at 1/3 the average county wage rate.

CV_j	The coefficient of variation for monthly average reservoir levels at site j.
SUB_i	A substitute index for county i equal to:
	$[\Sigma (SIZE_k/DISTANCE_k)]$ for k=1,...,m where m is the number of lakes or reservoirs greater than 500 surface acres within 250 miles one-way driving distance from county i, $SIZE_k$ is the surface acres of lake k, and $DISTANCE_k$ is the travel distance from county i to site k.
$SURACRES_j$	The surface acres of reservoir j.
$PICNIC_j$	The number of day-use picnic tables at site j.
$PARKING_j$	The number of day-use parking spots at site j.
$LANES_j$	The number of boat launch lanes at site j.

The model estimated is:

$$VISITS_{ij}=\exp(\beta_0)*POP^{\beta_1}*TOTCOST^{\beta_2}*LOGSUB^{\beta_3}*CV^{\beta_4}*SURACRES^{\beta_5}*PICNIC^{\beta_6}*PARKING^{\beta_7}*LANES^{\beta_8}*\exp(\epsilon) \tag{3}$$

where ε is assumed normally distributed with a mean of 0 and a constant variance. This model is estimated in double-log form by taking natural logs of all variables. The estimated coefficients are elasticities and ε satisfies the OLS assumptions.

The auxiliary R^2 is measured by separate regressions of each independent variable on the remaining independent variables. One widely-used rule of thumb is that multicollinearity is problematic if any of these auxiliary R^{2}'s are greater than the R^2 of the full model. (Greene, 1993, p. 269). The high auxiliary R^{2}'s associated with the facility variables suggest that multicollinearity is a problem. Also, demand theory supports coefficients on the facility variables between 0 and 1 (positive but decreasing marginal returns from additional facilities). The negative sign on PARKING is troublesome as well as the implausibly high coefficient on SURACRES. Policy analysis based on these coefficients may be misleading because estimated impacts on visitation and benefits from changes in facilities are of the wrong sign and/or magnitude.

IV. THE EXPERIMENT

Two sets of Monte Carlo simulations are conducted. For the first case, all four facility coefficients are set equal at 0.50. In the second case, the coefficients differ, β_5=0.80, β_6=0.45, β_7=0.65, and β_8=0.30.

Each Monte Carlo simulation produces 100 random values of the error term, in which there is a known constant error variance, σ^2. Three values of σ^2 were selected for this experiment. Setting σ^2 at 1.6 produces the approximate percent of variance explained by the model (R^2) as obtained by estimating the model with real data on the dependent variable. The other values of σ^2 produce R^2's which represent the range of typical values obtained in published travel cost models. Setting σ^2 at 2.8 gives a mean R^2 of about 0.40 while R^2 is 0.20 when σ^2 is 4.7. Experiments are conducted with these different error variances to compare the performance of methods for dealing with collinearity with changes in the error variance. The three error variances and two sets of coefficients produce six total experiments using each of the four methods for dealing with collinearity described above.

The first approach for dealing with multicollinearity is to include all four facility variables, along with the non-facility variables in an OLS regression. This technique is expected to give unbiased estimates, but with large variances.

The second approach is to include only one of the facility variables, SURACRES, as a proxy for the extent of facility developments at a reservoir. For this method, the other three facility variables are dropped from the model to reduce multicollinearity. This method is widely known to produce biased estimates when the true coefficients on the dropped variables equal zero.

The third method is principal components analysis (PCA), a commonly described method for dealing with multicollinearity. PCA should produce biased coefficients with lower variances than the OLS model. With PCA, one must choose how many characteristic vectors to include (e.g., Green, p. 272). For this paper, one, two, and three characteristic vectors are used. If all four characteristic vectors are included, PCA produces the same results as OLS with the full set of regressors.

The fourth approach, one which we believe to be unique, is the construction of a facility index. The index is constructed by estimating four first-stage regressions with all four non-facility variables included along with one facility variable. Define the vector of upwards biased first-stage coefficients on the facility variables as $(\hat{\alpha}_{SURACRES}, \hat{\alpha}_{PICNIC}, \hat{\alpha}_{PARKING}, \hat{\alpha}_{LANES})$. The four collinear variables are aggregated into the following index.

$$INDEX = (SURACRES^{\alpha_{SURACRES}}) * (PICNIC^{\alpha_{PICNIC}}) * (PARKING^{\alpha_{PARKING}}) * (LANES^{\alpha_{LA}} \tag{4}$$

The variable INDEX defined in (4) is then specified in a second-stage OLS model with all non-facility variables. Because the model is Cobb-Douglas in structure, the estimated coefficient obtained by using the single variable INDEX, $\hat{\delta}$ can then be multiplied by the $\hat{\alpha}$s to obtain the final estimates for the facility variables. These final estimated are:

$$\hat{B}_5 = \hat{\alpha}_{SURACRES} * \hat{\delta}$$
$$\hat{B}_6 = \hat{\alpha}_{PICNIC} * \hat{\delta}$$
$$\hat{B}_7 = \hat{\alpha}_{PARKING} * \hat{\delta}$$
$$\hat{B}_8 = \hat{\alpha}_{LANES} * \hat{\delta}$$

Results obtained by using each of these four methods are compared using the criterion of mean square error.

V. RESULTS

Results are presented in Table 2. In general, they show that PCA and the index method developed for this paper produce lower MSE's than OLS with the full set of regressors. While OLS is unbiased, the variances on the coefficients are large.

Using only SURACRES produce a coefficient which is biased upwards considerably. For a coefficient on SURACRES known to be 0.80, the estimate averages about 1.6. When the coefficient is set at 0.50, an average estimate of 1.3 is produced. These results suggest that policy analysis will produce an economically inefficient allocation of water and other facilities, if based on estimates produced by dropping facility variables.

The index performed better when all coefficients were set equal at 0.50. In all six simulation experiments, the coefficients on LANES was overestimated. This bias contributed to most of the total MSE when the coefficients varied. When the coefficients were set equal, the index outperformed PCA in every case. However, when the coefficients are different, use of PCA with one characteristic vector generally produced lower MSE's than the index. The index method seemed to perform better, in comparison to PCA, as the value of σ^2 was increased. Table 2 reports on results for $\sigma^2 = 2.8$.

As more characteristic vectors are included, PCA produces higher MSE's. More characteristic vectors decreased the bias but increased the variances. Thus, these results suggest the use of only one characteristic vector in PCA. When σ^2=1.6 and the coefficients were different, PCA had a lower total MSE than the index method.

VI. CONCLUSIONS

Results presented in this paper have implications for natural resource policy when estimating travel cost demand models in which facilities are highly correlated. Results of the simulation experiments show that the two most common approaches to dealing with serious multi-collinearity, using the full set of regressors and dropping variables, both perform poorly according the MSE criterion. PCA produced lower

MSE's than these two approaches. The construction of a facility index was described and it performed well. In most cases, the index method produced lower MSE's than PCA. Further analysis of the index approach could further illustrate the situations for which it is applicable.

Table 1. OLS Regression Results with Actual Data and All Variables

Variable	Estimate	T-Statistic	Auxiliary R^2
Intercept	136.09	5.53	----
POP	0.95	12.77	0.13
TOTCOST	-3.97	-21.52	0.12
LOGSUB	-24.22	-5.46	0.29
CV	-0.80	-7.88	0.29
SURACRES	2.32	7.56	0.80
LANES	0.65	2.40	0.44
PARKING	-2.44	-6.05	0.84
PICNIC	0.56	3.31	0.58

$R^2 = 0.671$
F-Statistic = 86.500
n = 348

Table 2. MSE Simulation Results, $\sigma^2 = 2.8$

A. Coefficients Different

METHOD	SURACRES	PICNIC	PARKING	LANES	Total MSE
OLS (All variables)	0.314	0.242	0.702	0.506	1.766
OLS (SURACRES only)	0.931	--	--	--	--
PCA (1 c.v.)	0.098	0.109	0.084	0.200	0.492
PCA (2 c.v.'s)	0.102	0.245	0.103	0.410	0.862
PCA (3 c.v.'s)	0.169	0.228	0.169	0.483	1.050
Facility Index	0.057	0.013	0.021	0.388	0.481

B. Coefficients Equal

METHOD	SURACRES	PICNIC	PARKING	LANES	Total MSE
OLS (All variables)	0.432	0.150	0.796	0.341	1.719
OLS (SURACRES only)	1.568	--	--	--	--
PCA (1 c.v.)	0.083	0.062	0.100	0.060	0.307
PCA (2 c.v.'s)	0.091	0.155	0.110	0.220	0.577
PCA (3 c.v.'s)	0.211	0.141	0.166	0.323	0.842
Facility Index	0.013	0.037	0.049	0.050	0.150

REFERENCES

Greene, William H. Econometric Analysis, 2nd edition. MacMillian Publishing Company, New York. 1993.

Judge, G., C. Hill., W. Griffiths, and T. Lee. The Theory and Practice of Econometrics, 2nd edition. John Wiley and Sons, New York. 1985.

Kling, Catherine L. "A Simulation Approach to Comparing Multiple Site Recreation Demand Models Using Chesapeake Bay Survey Data". Marine Resource Economics, vol. 4:95-109, 1987.

Mendenhall, William, Dennis D. Wackerly, and Richard L. Scheaffer. Mathematical Statistics with Applications, 4th edition. PWS-KENT Publishing Co., Boston. 1990.

Rosenthal, Donald H. "The Necessity for Substitute Prices in Recreation Demand Analyses". American Journal of Agricultural Economics, vol. 69(4):828-837, November 1987.

Smith, V. Kerry, William H. Desvouges, and Matthew P. McGivney. "Estimating Water Quality Benefits: An Econometric Analysis". Southern Economic Journal, vol. 50(2):422-437, October 1983.

Ward, Frank A. and John B. Loomis. "The Travel Cost Demand Model as an Environmental Policy Assessment Tool: A Review of the Literature". Western Journal of Agricultural Economics, vol. 11(2):164-178, December 1986.

Wetzstein, Michael E. and Richard D. Green. "Use of Principal Component Attractiveness Indexes in Recreation Demand Functions". Western Journal of Agricultural Economics, p. 11-21, July 1978.

THE IMPACTS OF RECREATIONAL SPENDING TO A LOCAL AND REGIONAL ECONOMY IN NORTHEASTERN ALABAMA

R. Jeff Teasley[1]

John C. Bergstrom[2]

Abstract

Effects of spending by people engaged in recreational activities can have profound impacts on local economies. Quantitative estimates of these impacts are therefore of interest to both recreation resource managers and decision-makers in local economies affected by these expenditures. An economic impact and interdependence analysis was conducted for expenditures associated with recreation at Lake Guntersville, a Tennessee Valley Authority (TVA) reservoir in Northeastern Alabama. The analysis considered five different aquatic plant coverage management alternatives. The impacts to the local economy of predicted recreation trips under each alternative were estimated using a trip response model linked to an input-output model. Implications of the results for future reservoir management strategies are discussed.

[1]Research Coordinator, Dept. of Agricultural & Applied Economics, University of Georgia.

[2]Associate Professor, Dept. of Agricultural & Applied Economics, University of Georgia.

THE IMPACTS OF RECREATIONAL SPENDING TO A LOCAL AND REGIONAL ECONOMY IN NORTHEASTERN ALABAMA

INTRODUCTION

Effects of spending by people engaged in recreational activities can have profound impacts on both local and regional economies. In some instances local economies are almost entirely dependant on recreational related dollars for their existence (e.g., mountain, tourist-based towns and villages). Quantifying the effect recreational spending has on a local or regional economy can be quite challenging. Surveying of an area's recreation users is the first step in the undertaking.

The focus of this paper will be the presentation of an economic impact and interdependence analysis that was completed on one of the Tennessee Valley Authority (TVA) reservoirs in Northeastern Alabama, Lake Guntersville. Lake Guntersville is located approximately thirty miles southeast of Huntsville, Alabama. It is one of the larger reservoirs in the TVA system with 67,900 surface acres and 949 miles of shoreline. Lake Guntersville has become renowned for its bass fishing opportunities as well as its support of myriad water-based activities. The economic impact and interdependence results were part of a larger effort which included use estimation procedures, user satisfaction and preference analyses, as well as valuation estimates based on CVM analyses.

The study was initiated by the TVA and US Corps of Engineers (USACE) as a result of high levels of exotic aquatic plant infestations in Lake Guntersville. High levels of aquatic plants result in increased problems associated with water-based recreation (e.g., fouled boat props, swimming area closings, residential shoreline access problems, decreased skiing opportunities, potential for large mosquito populations, etc.). However, aquatic plants also provide excellent fishing opportunities as well as excellent waterfowl habitats. Thus, there is a polarization of user group preferences with respect to plant coverage levels, which is of keen interest to the TVA and USACE from a management policy-making perspective.

Spinyleaf Naid, Eurasian Watermilfoil, and Hydrilla were the species of exotic aquatic plants which were at issue in the study. Watermilfoil is the leader of the three in terms of amount of coverage and has existed the longest in Lake Guntersville. The naid, least represented in amount of coverage, is also the least problematic of the exotic species. Hydrilla, though behind watermilfoil in amount of coverage, represents the highest potential problem plant of the three. Growth rate of hydrilla has been recorded at up to one inch per day under ideal growing conditions (i.e.- low flow rates, clear water, and high sunshine levels). The plant has also been found growing at depths to 18 feet. The combination of high growth rate potential and ability to grow at depth can spell trouble for lake managers.

The following analysis is presented in light of five different aquatic plant coverage management alternatives. For each alternative, the impacts to the local economy will be presented. Implications drawn from the results of the analysis are discussed last.

BACKGROUND CONCEPTS

Calculation of the economic interdependence and impact of recreation requires the estimation of the direct, indirect, induced, and total effects of increases in visitor expenditures (i.e., changes in final demand). These effects were estimated with IMPLAN[1] through the application of input-output analysis. Because of the importance of these concepts, direct, indirect, and induced effects are reviewed in this section.

Firms in a regional economy are economically interdependent. This interdependency takes the form of purchases of goods and services between firms which are needed for the provision of other goods and services for final delivery to consumers. Consider a regional economy surrounding a reservoir such as Lake Guntersville. Many businesses in this economy would provide goods and services to people visiting the reservoir for recreation and tourism. For example, a resort on the reservoir may offer a variety of goods and services to visitors including lodging, meals, supplies, boat rentals, etc. In order to provide these goods and services, the resort would have to purchase inputs (e.g.,

---

[1] IMPLAN refers to the input-output data base and modeling system developed by the U.S. Forest Service (Alward et al).

building and hardware supplies, fuel, food, professional services) from other businesses. These businesses, in turn, would have to purchase inputs from other businesses in order to increase their own output of goods and services.

Because of the interdependency between businesses, a change in consumer demand for the goods and services provided by the resort will impact other businesses in the economy. Suppose increased recreational visits to the region by people outside the region caused demand for lake resort goods and services to increase. The resort must purchase increased inputs from other businesses in order to meet this increased demand. For example, in order to provide more meals to visitors, the resort would have to purchase more food from businesses in the food production and distribution sectors. In order to increase output, businesses in the food production and distribution sectors would have to purchase more inputs from their suppliers. These purchases, in turn, would result in even more economic activity since the suppliers of inputs to the food production and distribution sectors would have to increase their purchases of inputs. Thus, the direct purchases of inputs by the resort in order to meet increased demand for goods and services by visitors initiates a "chain reaction" of additional purchases in the regional economy. The additional indirect purchases are the INDIRECT EFFECTS of the increase in demand for resort goods and services.

The direct and indirect effects of the increase in demand for resort goods and services can result in a overall increase in the production of goods and services in the regional economy. Such an increase in economic activity results in increased employment and household income. With the increase in household income, consumer demand for all types of goods and services could be expected to increase. This increase in consumer demand would stimulate further purchases of inputs in the regional economy. These purchases, which result from increased consumer expenditures, are the INDUCED EFFECTS of the original increase in demand for resort goods and services.

In summary, input-output analysis in general and IMPLAN in particular measure the total economic impact on the region of a net increase in consumer demand for goods and services produced by businesses within the region. Total economic impact is composed of direct, indirect, and induced effects of the change in consumer demand. Direct effects are the "first round" purchases of inputs by the firm experiencing the increase in demand. Indirect effects are purchasers made in the "second round", "third round", "fourth round", and so forth in order to support the production of input purchases in the "first round". Induced effects are increased purchases of goods and services in the region which result from increased income to households and owners of firms.

METHODOLOGY

Survey Design

Because of the large size of the reservoir and the complex nature of the study, a intensive sampling plan was developed to capture the users of Lake Guntersville. Over 80 interviewing sites were identified along the 949 shoreline miles of the lake and surveys were delivered in a face-to-face format. Much socioeconomic, trip, preference, and demographic information was collected via the on-site longform survey administered.

Expenditure data was collected via a mail-back survey handed out to respondents after the face-to-face survey was completed. The survey was organized into two parts: the first, collected recreation equipment expenditures on an annual basis; the second, collected trip expenses from the users for their current trip to the site. Each of the sections were grouped according to major types of expenses (e.g., boat equipment expenses, fishing equipment expenses, off-road vehicle expenses, food expenses, clothing expenses, etc.). Each major expenditure section then gained further detail by asking for specific information concerning the major type of expense (e.g., boat trailer tires, boat accessories, rods and reels, drinks, processed food, shoes, apparel, etc.).

After being handed the survey, respondents were asked to complete the survey once they returned home and return it in the self-addressed, postage paid envelope included in the package. If, after two weeks of the initial survey date, there was no response from the respondent a reminder postcard was sent urging the respondent to complete and return the survey. Two weeks following the postcard mailing, if no response was recorded the respondent was sent another survey, postage-paid envelope, and cover letter again urging them to complete the information and return it in the provided envelope.

695 expenditures were finally collected which represented a 39.8 percent return rate. Of the 695 surveys received, 673 were filled out sufficiently to allow their use in the estimation procedures. The modified Dillman Total Design Method of reminder scheduling for mail surveys performed fairly well in this case.

Visitation Changes

An extensive on-site sampling plan/use estimation procedure was initiated on Lake Guntersville. The result of this procedure produced use estimates which were used as baseline estimates for the study (see Table 1). On-site interviewing began during early 1991 and ended in early 1992. Aquatic plant coverage during this year was very close to 8,000 acres, or 10% total lake coverage.

Calculation of visitation change between five management alternatives was undertaken to assess the sensitivity of aquatic plant coverage change on visitation. In a separate survey created for this purpose, respondents were asked how many trips they would make to Lake Guntersville given each of the management scenarios presented.

Using the Trip Response Method (Teasley and Bergstrom) an equation was estimated of the general form:

(3.1) TRIPS $= f($PLNLIKE, NUMPEEP, BINCOME, RURAL, DIST, SEX, RESIDENT, AGE, M1, M2, M3, M4$)$,

where TRIPS is the estimated number of respondent trips given a specific plant coverage level, PLNLIKE is how much aquatic plant coverage respondents would like to see (1 being a small amount, 5 a large amount), NUMPEEP is number of people in household, BINCOME is the household income, RURAL is a dummy variable representing whether or not the respondent's household was located in a rural or urban area, DIST is the distance from the respondent's home to Lake Guntersville, SEX is the sex of the respondent, RESIDENT is whether or not the respondent lives on the lake (also a dummy variable), AGE is respondent's age, and M1 - M4 are dummy variables representing the different management alternatives. The TRM equation produced estimates of annual number of trips for an individual respondent under each management alternative.

Impact analysis is concerned with changes in the number of visitors from outside the impact region (OA visitors), not residents of the impact region (IA visitors). For this reason, it was necessary to compute increases, or decreases, in OA visitation for each management alternative. It was assumed that changes from the baseline management alternative to any management alternative would not alter the proportion of OA visitation for any user type. The total increase, or decrease, in number of visits for each user type and management alternative was calculated by multiplying the baseline number of visits by the appropriate percentage increase (decrease) of visits. Multiplying the total visitation increase by the proportion of on-site respondents who were OA visitors for that user type yielded the increase (decrease) in OA visitors necessary for the impact analysis. Proportions of IA/OA users for the impact area are presented in Table 2. For estimated number of visits by user type and management alternative see Table 3.

On-Site Visitors Impact Analysis

The impact of recreation on a local or regional economy is measured by the direct, indirect, and induced effects of spending associated with recreation trips. According to export base theory, growth in exports causes economic growth. Purchases of goods by OA visitors are exports, as they bring "new" dollars into the local region. Thus, impacts in an economy attributable to recreation are traceable to spending by OA visitors for recreation and related services.

This study estimated the impacts resulting from the five management alternatives as follows:

Minimum Mngt. Alternative	-	34,000 Acres aquatic plant coverage (Approx. 50% coverage - Hypothetical)
Mngt. Alternative A	-	20,200 Acres aquatic plant coverage (Approx. 30% coverage - 1988 Highest historical amount)
Mngt. Alternative B	-	14,200 Acres aquatic plant coverage (Approx. 20% coverage - 1989)

Mngt. Alternative C - 8,000 Acres aquatic plant coverage (Approx. 10% coverage - 1990)

Mngt. Alternative D - 0 Acres aquatic plant coverage.

Three of the five management scenarios discussed above were historically referenced to years preceding the surveying efforts of this project. Because of the extremely large proportion of users who had prior experience with Lake Guntersville (many for a decade or more), results of the analysis will reflect that experience. One important item in this experience, especially during high aquatic plant coverage years, was the plant control effort of TVA in priority areas around the lake (e.g., boat docks, boat ramps). The presentation of the management scenarios to survey respondents was consistent with the assumption that TVA would continue to control plants in priority areas under each level of coverage.

Only expenditures made in the impact area by OA users were relevant in determining final demand changes for impact analysis. Lake Guntersville lies in Marshall and Jackson counties, Alabama[2]. For this study, the impact area was chosen by taking the two counties in which the lake exists and those counties contiguous to them. The resulting eleven county impact area reached into three states, Alabama, Georgia, and Tennessee. The eleven counties are shown in Table 4.

All trip related expenses made on-site were assumed to occur in the impact area. OA user expenses incurred at home were not attributable to the impact area and were deleted. En route expenses were attributed to the impact area according to the percentage of OA users travel within the impact area. This percentage was determined by taking the proportion of straight-line distance from the respondent's home to the lake, to the amount of that straight line distance that lay within the impact area.

Annual equipment expenditures attributable to an impact area were estimated in the following manner. First, if the respondent was an OA user, all home expenses were deleted. Remaining expenditures were divided by the number of times the respondent reported they had visited the interview site or area with the equipment in order to put the annual dollar amount on a per trip basis. Mean

[2]It is noted that a small 'riverine' portion of Lake Guntersville extends into a part of Marion county, Tennessee.

expenditures per trip were then divided by the reported number of persons for whom the respondent paid to get an average expenditure profile per person per trip.

At the beginning of the analysis, expenditure profiles were estimated across four visitor types: Boater/Fisher, Boater/Non-fisher, Non-Boater/Fisher, and Non-boater/Non-Fisher. The mean expenditure profiles between these groups were not determined to be significantly different, however. Therefore, the user groupings were changed to Fishers/Non-Fishers and Boaters/Non-Boaters. Mean expenditures between Fishers/Non-Fishers and between Boaters/Non-Boaters were significantly different.

The two groupings, Fisher/Non-Fisher and Boater/Non-Boater are not mutually exclusive, however. We therefore chose to conduct our analysis using the Fisher/Non-Fisher grouping because the distinction between Fishers and Non-Fishers seemed of most interest to the study.

The final step required to prepare the expenditure profiles for I-O analysis was to allocate the expenditures by item across economic sectors to derive final demand effects by sector. The procedures involved using national annual personal consumption data (PCE) prepared by the Bureau of Economic Analysis (BEA) to develop percentage allocations to each of the appropriate sectors for each expenditure item (Watson and Brachter).

On-Site Visitor Interdependence Analysis

A related method for assessing the importance of recreation to a local area is to examine the total amount of business transactions in the economy generated by recreation visitation. This analysis resembles impact analysis, but considers the total effects to the economy of the recreation resource by including the recreation spending of IA visitors in the calculations. The procedures are similar to those outlined above, but incorporate slight modifications to the expenditure and visitation profiles.

Since interdependence analysis includes both IA and OA visitors, visitation levels are increased over those found in impact analysis. This, correspondingly, increases dollar value estimates of total effects. Mean expenditure profiles for OA visitors remain unchanged. Expenditure profiles for IA visitors include at home, en-route, and on-site expenses.

It is important to note that economic interdependence estimates do not measure the contribution of recreation spending to economic growth. The economic interdependence estimates simply show the amount of economic activity that is in some way related to reservoir recreation. In order to measure economic growth and development potential (which is often of primary policy and planning interest), economic impact estimates must be used -- not economic interdependence estimates.

ECONOMIC IMPACT AND INTERDEPENDENCE RESULTS FOR ON-SITE VISITORS

Expenditure Profiles

The economic impacts of recreation to an area depend on the amount of money OA visitors spend in the area. Tables 5 and 7 show the mean trip expenditures per person, per trip for OA visitors and for IA and OA visitors to the 11-county impact region. These expenditures are categorized by major trip expenditure type. Tables 6 and 8 represent mean equipment expenditures per person, per trip for OA visitors and for IA and OA visitors of the impact region. All dollar amounts are reported in 1990 dollars.

OA visitors spent more on average in every expenditure category than IA visitors. This is expected since they must travel a longer distance and therefore spend more in travel and perhaps lodging. Fishers and Boaters spent more than those who did not fish or boat. This is also expected as the expense of operating and traveling with the equipment needed to boat and/or fish requires more expense than activities other than boating or fishing.

The expenditures made by visitors for specific items are allocated to specific economic sectors, including those representing the primary manufacture of the goods purchased. This allocation process is termed margining. Since specific sectors are perturbed, one must allocate a specific percentage of the total respondent expenditure to each sector. The example follows through for all of the recreation expenditures attained in the expenditure survey. This allocation method was designed for compatibility with the economic sector definitions in IMPLAN.

Impact Results

Economic impact results are reported for total gross output, total income, and total employment categories. Total gross output is the sum of all annual industry sales or the annual value of outputs produced by industries in an economy. Total income is the sum of employee compensation (wages and salaries paid to employees of industries in a economy) and property type income. Property type income is defined as profits, rents, royalties, etc., paid to owners of property and firms that are engaged in the production of outputs in an economy (Palmer and Siverts). Estimated economic impacts (total effects) include the direct, indirect, and induced effects of recreational spending by OA users.

The general formula for attaching economic impact and interdependence effect dollar amounts to each alternative for each user group was specified as follows:

$$(3.1) \qquad DT = [(V^{S}_{(F,NF)} * A^{CNTY}_{(F,NF)}) * UI^{CNTY}_{(F,NF)}] \,/\, 1000.$$

DT equals dollar total impact(effect), $V^{S}_{(F,NF)}$ equals visitation level under the given management alternative for either Fishers or Non-Fishers, $A^{CNTY}_{(F,NF)}$ equals the percentage of users who did not live in the given impact area for either Fishers or Non-Fishers (OA users), and $UI^{CNTY}_{(F,NF)}$ equals the dollar impact(effect) per 1000 visits for either Fishers or Non-Fishers to a specific impact area.

Economic impacts for the 11-county impact area across each of the five management scenarios are shown in Tables 9, 10, and 11. Impacts for Fishers and Non-Fishers are shown broken out of the total impacts.

Because Fishers prefer more to less plants, a definite trend develops as plant coverage levels decrease. The exception is the minimum management alternative where, evidently, the users felt that the coverage levels would be excessive to the point as to prevent, or hinder, their main activity.

Non-Fishers have a mirror-image trend of the Fishers. Their preferences are for less coverage but kinks where the management alternative removes all plant coverage. This is perhaps in recognition to the biological necessity of aquatic plants in the lake ecosystem.

Interdependence Results

Interdependence results for the 11-county local impact area are presented in Tables 12, 13, and 14[3]. Total dollar figures for the interdependence analysis are larger than for the impact analysis primarily because IA users are included. Clearly, the addition of these users adds a tremendous effect to the values presented. Again, it is important to note that these values represent <u>contribution</u> to the economy by recreation expenditures rather than <u>impacts</u>.

Residential Economic Contribution

Homeowners[4] of the lake were also surveyed. Residential sampling produced approximately 370 usable expenditure sections for analysis. The values obtained from this analysis are reported separately because of the separate use estimation procedures applied between this group and the on-site surveying. Also, residents were conjectured to have different expenditure patterns than on-site users. Trip expenses for residents were considered to be inseparable from the daily functioning of the household and were therefore not used in IMPLAN calculations. Annual equipment use and expenditures were, however, used for the IMPLAN calculations.

Use estimates for residents are presented in Table 15. Estimated number of visits for residents by user group and management alternative are presented in Table 16. Because the IMPLAN results for residents are of an interdependence type, no visitation adjustment was needed for IA and OA users as was done for <u>impact</u> analysis. Residential equipment expenditure profiles are presented in Table 17 by user group. These mean expenditures are on a per household basis.

Interdependence effects for the 11-county economy by residents are described in Tables 18, 19, and 20 by user group and total. Residential expenditure contributions to the economy were measured by

---

[3] These figures do not include expenditures influencing the economy by people living <u>directly</u> on the lake.

[4]Eligible households, for the purposes of our study, included all those properties which were directly on the lake and also those which were off the water but still retained legal access.

allocating residential expenditures across the appropriate IMPLAN sectors as performed with the previous impact and interdependence analyses.

Residential and On-Site Visitor Contribution to Local Economy

In order to get a more complete picture of the contribution Lake Guntersville recreation users make to the 11-county impact area the on-site and residential interdependence analyses were summed. The combined residential and on-site visitor economic contributions to the impact area are presented below by user group and management alternative (see Tables 21, 22, and 23).

IMPLICATIONS

Public reservoirs in the United States such as those managed by the Tennessee Valley Authority and the U.S. Army Corps of Engineers provide multiple services to society. These services include flood control, navigation, hydropower, recreational opportunities, and environmental amenities (e.g., aesthetic enjoyment by visitors and lake residents). Achieving a reasonable and publicly-acceptable balance between the many, often competing, uses of public reservoirs is a major management challenge. The management of public reservoirs can be facilitated by information on the economic value and impact of reservoir management strategies.

Management of non-native aquatic plants in public reservoirs is of considerable interest and concern in the United States, particularly in the Southeast. In the absence of control efforts, with favorable growing conditions, non-native aquatic plant coverage can rapidly expand in a reservoir. For example, in Lake Guntersville, non-native aquatic plant coverage reached a high of about 30% of total reservoir acreage in 1988 in spite of heavy control efforts.

Large amounts of non-native aquatic plant coverage can have both negative and positive effects on reservoir uses and services. Negative effects include clogging water intake pipes, restricting commercial navigation, restricting access to boat docks and boat launching areas, promoting mosquito production, and interfering with certain recreational/leisure experiences (e.g., recreational boating, water skiing, swimming). Positive effects of aquatic plants include provision of food, cover, and oxygen for waterfowl, fish, and other wildlife species, and enhancement of certain recreational/leisure experiences (e.g., fishing, waterfowl hunting, wildlife observation). This study focused on examining relationships between non-native aquatic plant management and reservoir-based recreation at Lake Guntersville, Alabama.

Over the years, reactions to the amount of aquatic plant coverage in Lake Guntersville from recreational/leisure users of the lake has been mixed. Recreational boaters, for example, have complained about plants getting tangled in boat motor propellers, and recreational swimmers have complained about aquatic plants interfering with lake swimming opportunities. On the other hand, in the past many fishers have attributed good fishing opportunities at Lake Guntersville to the presence of non-native aquatic plants. There has also been concern expressed in the past that aquatic plant management may have major effects on the economic value and impacts of recreational/leisure uses of Lake Guntersville. Before this study, however, very little quantitative information has been available to assess lake user's preferences and attitudes with respect to aquatic plant coverage and the effects of changes in aquatic plant coverage on economic value and impacts.

Five alternative plant coverage scenarios were evaluated in this study. Under the "Minimum Control" Alternative, 34,000 acres of the lake (about 50% of the total reservoir surface area) would have aquatic plants "topped out" or just below the water surface. Under "Management Alternative A", 20,200 acres of the lake (about 30% of the total reservoir surface area) would have aquatic plants "topped out" or just below the water surface. Under "Management Alternative B", 14,200 acres of the lake (about 20% of the total reservoir surface area) would have aquatic plants "topped out" or just below the water surface. Under "Management Alternative C", 8,000 acres of the lake (about 10% of the total reservoir surface area) would have aquatic plants "topped out" or just below the water surface. Under "Management Alternative D", aquatic plant coverage would be reduced to near zero acres (representing, for example, an elimination of non-native aquatic plants in the lake).

The economic analysis included estimates of economic impact and interdependence. The estimates of economic impact answer the question, "How much does recreational use of Lake Guntersville contribute to the economic growth of the local economy around the lake?" Economic growth results from the influx of "new money" into the local economy as visitors who live outside of the local economy (non-residents) spend money on their visits to the lake. Thus, the economic impact results only include the effects of non-resident expenditures.

The greatest amount of economic growth was projected to result from "Management Alternative A" and "Management Alternative B". Both of these alternatives would result in about $160 million worth of total gross output being added to the 11-county local region surrounding Lake Guntersville. The smallest amount of economic growth was projected to result from the "Minimum Control" Alternative and "Management Alternative D". These alternatives would result in about $120 million worth of total gross output being added to the 11-county region surrounding Lake Guntersville. "Management Alternative C" would contribute about $150 million worth of total gross output to the 11-county region.

The economic impact estimates provide a measure of the additional economic activity (e.g., output, income, jobs) in a regional economy which are directly attributable to Lake Guntersville. If reservoir management or changes in visitation patterns result in a reallocation of trips by non-residents away from Lake Guntersville, the resulting decrease in recreation expenditures and economic activity would represent a loss to the local economy surrounding the lake. Alternatively, if reservoir management or changes in visitation patterns resulted in increased trips from non-residents to Lake Guntersville, the resulting increase in recreation expenditures and economy activity would represent a gain to the local economy.

The estimates of economic interdependence answer the question, "How much economic activity in a region is in some way tied to recreation at Lake Guntersville?" Economic interdependence considers the effects of all reservoir-related expenditures, including expenditures by non-resident visitors and lake users who live within the boundaries of the local economy surrounding the lake (e.g., households located on or near the lake). These expenditures by non-resident visitors and resident users contribute to both economic growth and stability in the local economy.

Economic interdependence effects were greatest for "Management Alternative A" and "Management Alternative B". Under both of these alternatives, about $580 million worth of total gross output in the 11-county region surrounding Lake Guntersville is projected to be associated with lake-related expenditures. Economic interdependence effects were smallest for the "Minimum Management" Alternative and "Management Alternative D". Under both of these alternatives, about $420 million worth of total gross output in the 11-county region is projected to be associated with lake-related expenditures. About $518 million of total gross output is projected to be associated with lake-related expenditures under "Management Alternative C".

The economic interdependence effects indicated the interdependency of recreation with economic activity in the local economy surrounding Lake Guntersville. If recreation is highly interdependent with businesses and other industries in the local economy, recreational expenditures will tend to contribute substantially to the magnitude and distribution of goods and services output, income, employment, and other economic dimensions of the economy. Thus, if recreation use of the reservoir were to drastically decrease, there may be a large "shock" to the local economy. The effects of this shock would eventually diminish as residents reallocate expenditures to other businesses and industries within and outside of the local economy. However, a certain amount would be lost as non-resident dollars stopped flowing into the local economy. Also, if residents do not reallocate expenditures within the local economy, the local economy may suffer a net loss in economic activity such as income and employment. This would occur, for example, if residents and non-residents reallocated their recreation expenditures to recreational areas outside of the local economy surrounding Lake Guntersville.

These results suggest that there may by large resulting economic effects associated with differing levels of aquatic plants in reservoirs. Because large reservoirs, such as Guntersville, support varied recreational activities that are affected in different ways by aquatic vegetation control practices of the plants could have potentially large effects on local economies. Economic impact and interdependence information, such as is presented here, is important input into the management decision-making processes, as indicated by the demand for such information by natural resource managers, local business operators, and other decision-makers.

REFERENCES

Alward, G.S., H.G. Davis, K.A. Despotakis, and E.M. Lofting. "Regional Non-Survey Input-Output Analysis with IMPLAN." Presented paper, Annual meetings of the Southern Regional Science Association meetings, Washington, DC, 1985.

Dillman, D. A. Mail and Telephone Surveys. New York: John Wiley & Sons, 1978.

Findeis, J.L. "Input-output Analysis: A Methodology for Assessing Short Run Changes in Exogenous Prices, Output, and Final Demand." Department of Agricultural Economics and Rural sociology, The Pennsylvania State University, Publication #165, Spring, 1983.

Palmer, C. and E. Siverts. IMPLAN Analysis Guide. Land Management Planning Systems Section, U.S.D.A. Forest Service, Ft. Collins, Colorado, 1985.

Teasley, R. Jeff, and John C. Bergstrom. "Estimating Revenue Capture Potential Associated with Public Area Recreation." Forthcoming in the Journal of Agricultural and Resource Economics. July, 1994.

Watson, A.E. and L. Brachter. Public Area Recreation Visitor Study: Phase III Reporting. Final cooperative research agreement report to Southeastern Forest Experiment Station, Athens, Georgia. 1987.

Table 1. "Baseline" Use at Formal Sites			
User Group	Total On-site Rec Visits	Confidence Interval (90%)	% of Total Visits
All Users	2,844,718.14	29.63	100%
Fishers	976,508.47	24.33	34%
Non-Fishers	1,868,209.67	43.74	66%

Table 2. Percentage of Respondents in IA User/OA User Category for Each Impact Region		
11 County Impact Region	Fishers	Non-Fishers
IA users	74.3	72.6
OA users	25.7	27.4

Table 3. Estimated Number of Visits by User Group and Management Alternative					
User Group	Min Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	1,644,440.26	1,847,554.03	1,429,608.4	976,508.47	476,536.13
Non-Fisher	855,640.03	1,434,785.03	1,765,458.14	1,868,209.67	1,750,512.46

Table 4. Counties Used in Impact and Interdependence Analysis			
State	County	Population (thousands)	Area (Sq. Miles)
Alabama	Blount	39.2	646
	Cullman	67.6	738
	De Kalb	54.7	778
	Etowah	99.8	535
	Jackson	47.8	1079
	Madison	238.9	805
	Marshall	70.8	567
	Morgan	100	582
Georgia	Dade	13.1	174
Tennessee	Franklin	34.7	553
	Marion	24.9	500
Totals		791.7	6,957

Table 5. Mean Trip Expenses for OA Users of the 11-County Impact Region (Reported dollar amounts are per person, per trip, 1990 dollars)

User Type	Lodging ($)	Food ($)	Transpt ($)	Activities ($)	Misc. ($)	Total ($)
	Expenditure Category					
Fisher	52.88	43.21	77.27	4.85	6.90	185.11
Non-Fisher	8.33	58.90	67.30	4.95	22.12	161.60

Table 6. Mean Equipment Expenditures for OA Users of the 11-County Impact Region (Reported dollar amounts are on a per person, per trip basis, 1990 dollars),

Expenditure Category	User Type	
	Fisher	Non-Fisher
Motor Boat	43.16	78.61
Other Boats	0.06	2.86
Skiing	0.30	0
Camping Vehicles	5.14	2.51
Backpacking	0.41	0.18
Fishing	20.46	3.14
Bike	0	0.36
Motor Bike	0	0
Off-Road Vehicles	0.68	0
Hunting	2.98	0.02
All Other Equipment	0.51	0.06

Table 7. Mean Trip Expenses for IA Users and OA Users of the 11-County Impact Region (Reported dollar amounts are per person, per trip, 1990 dollars)

User Type	Lodging ($)	Food ($)	Transpt ($)	Activities ($)	Misc. ($)	Total ($)
	Expenditure Category					
Fisher	21.48	35.70	67.36	7.20	22.32	154.06
Non-Fisher	5.78	49.60	51.87	4.41	14.90	126.56

Table 8. Mean Equipment Expenditures for IA Users of the 11-County Interdependence Region (Reported dollar amounts are on a per person, per trip basis, 1990 dollars)

Expenditure Category	User Type	
	Fisher	Non-Fisher
Motor Boat	80.00	63.16
Other Boats	1.89	2.77
Skiing	0.16	0.33
Camping Vehicles	5.16	11.45
Backpacking	0.48	2.21
Fishing	15.06	5.17
Bike	2.17	4.46
Motor Bike	1.48	3.69
Off-Road Vehicles	1.12	4.47
Hunting	4.65	19.38
All Other Equipment	0.44	1.50

Table 9.	Total Gross Output Due to OA Users Recreational Visits to Lake Guntersville Under Different Management Alternatives, 1990 Dollars				
User Type	Total Gross Output - Economic Impact (Million $)				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	75.61	84.95	65.73	44.90	21.91
Non-Fisher	48.91	82.01	100.91	106.78	100.05
Total	124.52	166.96	166.64	151.68	121.96

Table 10.	Total Income Due to OA Users Recreational Visits to Lake Guntersville Under Different Management Alternatives, 1990 Dollars				
User Type	Total Income - Economic Impact (Million $)				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	41.25	46.34	35.86	24.49	11.95
Non-Fisher	26.56	44.54	54.81	58.00	54.34
Total	67.81	90.88	90.67	82.49	66.29

Table 11.	Total Employment Due to OA Users Recreational Visits to Lake Guntersville Under Different Management Alternatives				
User Type	Total Employment - Total Number of Jobs				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	1,948.28	2,188.93	1,693.76	1,156.94	564.59
Non-Fisher	1,289.45	2,162.22	2,660.55	2,815.39	2,638.02
Total	3,237.73	4,351.15	4,354.33	3,972.33	3,202.61

Table 12. Total Gross Output Due to IA and OA User Recreational Visits to Lake Guntersville Under Different Management Alternatives, 1990 Dollars

User Type	Total Gross Output - Economic Impact (Million $)				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	277.09	311.31	240.89	164.54	80.30
Non-Fisher	162.06	271.75	334.38	353.84	331.55
Total	439.15	583.06	575.27	518.38	411.85

Table 13. Total Income Due to IA and OA User Recreational Visits to Lake Guntersville Under Different Management Alternatives, 1990 Dollars

User Type	Total Income - Economic Impact (Million $)				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	151.12	169.79	131.38	89.74	43.79
Non-Fisher	88.82	148.93	183.25	193.92	181.70
Total	239.94	318.72	314.63	283.66	225.49

Table 14. Total Employment Due to IA and OA User Recreational Visits to Lake Guntersville Under Different Management Alternatives

User Type	Total Employment - Total Number of Jobs				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	6,692.87	7,519.54	5,818.51	3,974.39	1,939.50
Non-Fisher	4,089.96	6,858.27	8,438.89	8,930.04	8,367.45
Total	10,782.83	14,577.81	14,257.40	12,904.43	10,306.95

Table 15. Residential Use Estimates in Visits by User Group			
User Group	Total On-site Rec Visits	Confidence Interval (90%)	% of Total Visits
All Resident Users	208,221.87	9.94	100%
Fishers	89,666.69	17.3	43%
Non-Fishers	107,708.80	14.88	52%

Table 16. Estimated Number of Visits by User Group and Management Alternative					
User Group	Min Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	150,998.71	169,649.38	131,272.03	89,666.69	43,847.01
Non-Fisher	49,330.63	82,612.65	101,784.82	107,708.80	100,815.44

Table 17. Residential Household Mean Annual Equipment Expenditure Profiles for Fishers and Non-Fishers, Per Average Trip, 1990 Dollars		
Expenditure Category	User Type	
	Fisher	Non-Fisher
Motor Boat	31.75	65.13
Other Boats	0.60	1.88
Skiing	0.0	0.48
Fishing	3.99	2.49
Hunting	3.73	1.36
All Other Equipment	0.09	1.03

Table 18.	Total Gross Output Due to Residential Recreational Spending at Lake Guntersville Under Different Management Alternatives, 1990 dollars				
User Type	Total Gross Output - Economic Impact (Million $)				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	2.36	2.65	2.05	1.40	.68
Non-Fisher	1.33	2.22	2.74	2.90	2.71
Total	3.69	4.87	4.79	4.30	3.39

Table 19.	Total Income Due to Residential Recreational Spending on Lake Guntersville Under Different Management Alternatives, 1990 dollars				
User Type	Total Income - Economic Impact (Million $)				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	1.34	1.51	1.17	.80	.39
Non-Fisher	.75	1.26	1.56	1.65	1.54
Total	2.09	2.77	2.73	2.45	1.93

Table 20.	Total Employment Due to Residential Recreational Spending on Lake Guntersville Under Different Management Alternatives				
User Type	Total Employment - Total Number of Jobs				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	57.38	64.47	49.88	34.07	16.66
Non-Fisher	30.58	51.22	63.11	66.78	62.51
Total	87.96	115.69	112.99	100.85	79.17

Table 21.	Total Gross Output Due to Residential and On-Site Recreational Spending at Lake Guntersville Under Different Management Alternatives, 1990 Dollars				
User Type	Total Gross Output - Economic Interdependence (Million $)				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	279.45	313.96	242.94	165.94	80.98
Non-Fisher	163.39	273.97	337.12	356.74	334.26
Total	442.84	587.93	580.06	522.68	415.24

Table 22.	Total Income Due to Residential and On-Site Recreational Spending on Lake Guntersville Under Different Management Alternatives, 1990 Dollars				
User Type	Total Income - Economic Interdependence (Million $)				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	152.46	171.30	132.55	90.54	44.18
Non-Fisher	89.57	150.19	184.81	195.57	183.24
Total	242.03	321.49	317.36	286.11	227.42

Table 23.	Total Employment Due to Residential and On-Site Recreational Spending on Lake Guntersville Under Different Management Alternatives				
User Type	Total Employment - Total Number of Jobs				
	Min. Mngt. Alternative	Mngt. Alternative A	Mngt. Alternative B	Mngt. Alternative C	Mngt. Alternative D
Fisher	6,750.25	7,584.01	5,868.39	4,008.46	1956.16
Non-Fisher	4,120.54	6,909.49	8,502.00	8,996.82	8429.96
Total	10,870.79	14,493.50	14,370.39	13,005.28	10386.12

CONTINGENT VALUATION AND REVEALED PREFERENCE METHODOLOGIES COMPARING THE ESTIMATES FOR QUASI-PUBLIC GOODS

Richard T. Carson
Nicholas E. Flores
Kerry M. Martin
Jennifer L. Wright

ABSTRACT

A comprehensive literature search provides 79 studies from which 541 comparisons of contingent valuation (CV) to revealed preference (RP) estimates are made. Summary statistics of the CV/RP ratios are provided for the complete dataset, a 5% trimmed dataset, and a weighted dataset that gives equal weight to each study rather than each CV/RP comparison. For the complete dataset, the sample mean CV/RP ratio is 0.88 with a 95% confidence interval [0.80-0.95] and a sample median of 0.74. For the trimmed and weighted data sets, the summary statistics are (0.77; [0.73-0.81]; 0.74) and (0.92; [0.81-1.02]; 0.94), respectively. The correlation coefficients between the CV and RP estimates for the three datasets are 0.40, 0.60, and 0.68, respectively. Non-parametric density estimates are provided, as well as the results of regressions of the observed CV/RP ratios on the basic RP technique used and the broad class of goods valued.

An earlier version of this paper was presented at the Annual U.S. Department of Agriculture W-133 Meeting on Benefits and Costs in Natural Resources Planning, Tucson, Arizona. Please send comments to Richard Carson, Department of Economics, University of California, San Diego, La Jolla, CA 92093.

CONTINGENT VALUATION AND REVEALED PREFERENCE METHODOLOGIES COMPARING THE ESTIMATES FOR QUASI-PUBLIC GOODS

1. INTRODUCTION

Beginning with Knetsch and Davis (1966), the comparison of contingent valuation (CV) estimates for government-provided, quasi-public goods with estimates obtained from revealed preference (RP) techniques, such as travel cost analysis and hedonic pricing, has played a key role in assessing the validity and reliability of the contingent valuation method. In their assessment of the contingent valuation method twenty years later, Cummings, Brookshire and Schulze (1986) placed considerable emphasis on comparing estimates from eight studies that used both contingent valuation and revealed preference techniques for similar quasi-public goods.[1] The assemblage of studies in Cummings, Brookshire, and Schulze (1986) emphasized the shift away from treating revealed preference techniques as the "truth," toward the realization that revealed preference estimates are random variables which are sensitive to details such as commodity definition, the functional form used in estimation, and other technique-specific assumptions such as the value of time and the number of sites in a travel cost study. As a result of this shift, comparisons between contingent valuation and revealed preference estimates are generally assumed to represent tests of convergent validity rather than criterion validity.[2] Still, such comparisons can play a prominent role in discussions of whether there is a need to "calibrate" contingent valuation estimates up (Hoehn and Randall, 1987) or down (Diamond and Hausman, forthcoming) and issues such as whether contingent valuation estimates systematically vary with the good being valued.

The focus of this paper is to summarize the available information and provide readers with the broadest possible overview of how CV estimates for quasi-public goods correspond with estimates based

---

[1]The eight studies Cummings, Brookshire, and Schulze (1986) looked at were Knetsch and Davis (1966), Bishop and Heberlein (1979), Thayer (1981), Brookshire *et al.* (1982), Desvousges, Smith and McGivney (1983), Sellar, Stoll and Chavas (1985), Brookshire *et al.* (1985), and Cummings *et al.* (1986).

[2]Tests of criterion validity are possible when one is comparing an estimate from some technique to a value which is known to be the truth. Tests of convergent validity are possible when there are two or more measurement techniques potentially capable of measuring the desired quantity, but both techniques do so with error. Mitchell and Carson (1989) provide a discussion of this issue.

on revealed preference techniques. Through an extensive search of both the published and unpublished literature, we located 79 studies that provide 541 comparisons of contingent valuation to revealed preference estimates.

2. STUDY INCLUSION CRITERIA

To help locate studies that contain both CV and RP estimates, we systematically reviewed entries in the Carson *et al.* (1994) bibliography of over 1600 contingent valuation papers. To be eligible for inclusion in our sample, a study must provide at least one contingent valuation estimate and one revealed preference estimate for essentially the same quasi-public good; thus, no studies of private goods are included. The goods valued are various forms of recreation (mostly outdoor), changes in environmental amenities such as air, noise, or water pollution, and changes in health risks. Consumers (individuals or households) had to have been interviewed to obtain the contingent valuation estimate. Thus, we did not include studies where the respondents were not consumers such as Bohm's (1984) study of willingness to purchase statistical information by local governments. Furthermore, we considered only contingent valuation estimates of willingness to pay (WTP). Therefore, we excluded estimates based on willingness to accept compensation or on contingent behavior responses.[3] Otherwise, we have tried to be inclusive with respect to study estimates.

The time spanned by the studies we examined is nearly thirty years, 1966-1994. The earliest study is Knetsch and Davis' (1966) well-known contingent valuation-travel cost comparison of outdoor recreation in Maine. The latest study considered is Choe, Whittington, and Lauria (1994) who value the opening of a polluted urban beach in the Philippines.

Due to well-known, potential biases in relying upon only the published literature to summarize research findings, we have spent considerable effort searching the unpublished literature including theses,

[3] We do include CV estimates derived from willingness to drive questions if they were intended to be directly compared to a travel cost estimate. CV questions phrased in terms of willingness to give up other goods are not included. No comparisons between CV willingness to pay estimates and actual willingness to accept compensation (*e.g.*, Bishop and Heberlein, 1979) are used as our initial investigation suggested that CV/RP ratios in such comparisons are almost always substantially below 1.0.

dissertations, conference papers, and government reports.[4] We have also drawn upon the rapidly growing nonmarket valuation literature from studies conducted outside the United States.

Multiple estimates from a single study are provided when the study valued multiple goods. This is common, for instance, in situations where respondents were interviewed at several recreational fishing locations and travel cost and contingent valuation estimates were made for each location (*e.g.*, Duffield and Allen, 1988) or where different levels of a good are valued (*e.g.*, Shechter, 1991). Multiple estimates are also provided when a study used different analytical assumptions (*e.g.*, Smith, Desvousges, and Fisher, 1986) in making the CV and/or RP estimates. In such cases, we considered all of the possible comparisons between the CV and RP estimates for the good in question. Studies often show a clear preference for a particular estimate and provide a rationale for the choice. However, the choice of a particular estimate is subjective, and when facing the same choices, different researchers might undoubtedly make different choices. To maintain as neutral of a position as possible, we considered all available comparisons made explicitly in the study or which are easily inferred.[5]

The studies considered provide value estimates for a wide variety of quasi-public goods. As a result, we look at everything from the value of a recreational fishing day on the Blue Mesa Reservoir in Colorado to the value of a statistical life estimated from national occupational risk data. There is a substantial amount of variation between the goods considered, between the changes in the goods valued, and between the specific implementations of the valuation techniques used. There is also variation both across and within studies and in how closely the goods in different CV and RP comparisons actually match-up. This variation is both a strength and a weakness. The variation allows for an analysis that favors a "big-picture" view, but at the same time may be of little relevance to a particular study. If there is a strong signal that CV, as a general valuation approach, substantially under- or over- estimates quasi-public goods'

---

[4]Berg (1994, p. 401) underscores this position based on his study of publication bias by noting that "If the meta-analysis is restricted to published studies, then there is a risk that it will lead to biased conclusions. This is especially problematic in that one of the major advantages of meta-analysis is that the aggregation of data can lead to effect size estimates with very small variance, giving the impression of conclusiveness in circumstances where the summary estimate is biased. That is, the resulting inferences may not only be wrong but appear convincing."

[5]We have strived to avoid including duplicate estimates or estimates obtained by simple transformations such as aggregation.

values relative to revealed preference techniques, one is likely to see it in a sample as large as ours. Small effects and subtle interactions between particular types of goods and very specific valuation techniques used may, however, be missed.

We coded the revealed preference techniques used in the papers into five broad categories. The first of these is single site travel cost models (TC1). The second is multiple site travel cost models (TC2). The third is hedonic pricing (HP). The fourth includes household production function, averting behavior, and expenditure function models (AVERT) not already included in TC2. The last category includes the creation of simulated or actual markets (ACTUAL) for the good. We excluded estimates from any technique which were not designed to capture net willingness to pay/consumer surplus such as actual trip expenditures. There are 241 TC1, 170 TC2, 52 HP, 26 AVERT, and 48 ACTUAL estimates.

We have also coded the goods valued in the various studies into three broad classes. The first class, recreation (REC), includes studies that valued outdoor recreation such as sport fishing, hunting, and camping. The second class is environmental amenities (ENVAM). Many studies in this class valued such goods as improved air and water quality. The third class is health risk (HEALTH). Studies in this class valued reductions in environmental health risks. There are 366 REC, 160 ENVAM and 15 HEALTH estimates. There is a considerable correspondence between the general class of good being valued and the RP technique used. This is particularly true of outdoor recreation where single (TC1) and multiple (TC2) site travel cost models are almost exclusively used.

3. BRIEF DESCRIPTION OF STUDIES

To help the reader, we have included a basic summary of each study used in our analysis. The study summaries are grouped into four categories based on their revealed preference methodology: travel cost (single-site and multiple-site)[6], hedonic pricing, averting behavior/household production functions, and simulated/actual market. Within each revealed preference methodology, the studies are organized

[6] Because studies sometimes provide both types of estimates, single- and multiple-site studies are combined.

chronologically. The number of CV/RP comparison ratios obtained from each study appears in parentheses after each summary.

3.1 Comparisons of Contingent Valuation with Travel Cost

Knetsch and Davis (1966) valued outdoor recreation at a forest recreation area in northern Maine using a single-site travel cost model and two variants of the contingent valuation approach. The data was obtained from on-site interviews of recreation area users. Both willingness to pay and willingness to drive additional distances were elicited from the respondents. (*2 comparison ratios*)

Beardsley (1970) conducted a study to value recreation on the Cache la Pandre River in Colorado using both travel cost and contingent valuation methodologies. The study data was obtained from an in-person survey of river visitors conducted in 1966. The authors present two different benefit estimates using a simple single-site zonal travel cost model and comparable contingent valuation estimates. (*2 comparison ratios*)

Binkley and Hanemann (1978) examined beach usage in Boston using both travel cost and contingent valuation methodologies. The data was obtained from in-person interviews of beach users. Using a multi-site travel cost model, the authors estimate a range of average values per day that can be compared to a contingent valuation estimate for the same. (*2 comparison ratios*)

In a study best known for its innovative elicitation of willingness to accept compensation in a simulated market, Bishop and Heberlein (1979) also used the contingent valuation and travel cost methodologies to value willingness to pay for goose hunting in the Horicon Zone of Wisconsin. The data was obtained from a mail questionnaire administered to a random sample of hunters who had applied for early season permits. Under different assumptions for the value of time, the authors present three travel cost estimates and one contingent valuation estimate. (*3 comparison ratios*)

Smith (1980) used both travel cost and contingent valuation models to estimate the recreational value at Oregon's Cullahy Lake in her Master's Thesis. The study data was obtained from on-site interviews of lake visitors conducted in 1979. The author presents one travel cost and one contingent valuation estimate. (*1 comparison ratio*)

Thayer (1981) used contingent valuation and a site-substitution travel cost model to estimate willingness to pay for the preservation of the Santa Fe National Forest (located in the Jemez Mountain Area of New Mexico) in its original state (*i.e.*, to prevent geothermal activity which was scheduled to begin in the early 1980's). The study data was obtained from interviews of recreationists conducted in the fall of 1976 and the spring of 1977. The authors present contingent valuation and travel cost estimates for daytrippers, campers, and the visitor population as a whole under different travel cost modeling assumptions. (*6 comparison ratios*)

Haspel and Johnson (1982) conducted a study to assess the impact of proposed surface mining to be located near Utah's Bryce Canyon National Park. A survey was administered to different samples of the park's visitors in the summer of 1980. One section of the survey was designed to generate data to allow for a comparison of travel cost and contingent valuation estimation techniques. Under alternate assumptions regarding model specification and the value of time, eight travel cost estimates are presented. Two contingent valuation estimates are provided that were calculated using the maximum additional distance visitors were willing to drive to visit Bryce Canyon. Johnson and Haspel (1983) used two additional survey samples to derive new travel cost estimates which were then compared to one of the CV estimates from Haspel and Johnson (1982). (*8, 2 comparison ratios*)

Vaughan and Russell (1982) valued a day of freshwater fishing by the species sought using the travel cost method and contingent valuation method. Their data was obtained from the 1975 National Survey of Hunting, Fishing and Wildlife Associated Recreation. Some of the parameters in the travel cost model were derived using data from a 1979 mail survey of recreational fee-fishing sites in the United States. The travel cost analysis used a varying-parameter model with multiple sites. For trout and catfish, the authors calculated two travel cost estimates, one without inclusion of time costs and one with time valued at the median wage by zone. A contingent valuation estimate was also provided for the two types of fish. (*4 comparison ratios*)

Desvousges, Smith and McGivney (1983) valued water quality improvements in the Monongahela river basin in Western Pennsylvania. In-person interviews were administered to a sample of area households. Three different scenarios were valued: avoiding a loss of water quality, an improvement from boatable to fishable water quality, and an improvement from fishable to swimmable water quality. One travel cost estimate for each scenario is presented along with four different contingent valuation estimates obtained from different elicitation methods. Smith, Desvousges and Fisher (1986) use the same CV data but present three new travel cost estimates for each scenario using different model specifications. (*12, 36 comparison ratios*)

Harris (1983) used both travel cost and contingent valuation methodologies to estimate the benefits from Colorado's fisheries in his Ph.D. dissertation. The study data was obtained from a mail questionnaire sent to a sample of Colorado fishing licensees. Both travel cost and contingent valuation estimates are presented for the full-sample and a single-purpose trip subsample by four fishery types (wild, basic yield, plains and combined). (*8 comparison ratios*)

Duffield (1984) conducted a study to estimate recreational values for the Kootenai Falls in northwestern Montana using the travel cost and contingent valuation methodologies. On site, in-person interviews were conducted in the summers of 1981 and 1982. Two travel cost estimates are presented under alternate assumptions regarding model specifications and contingent valuation estimates for two different payment vehicles. (*4 comparison ratios*)

ECO Northwest (1984) used simple travel cost, hedonic travel cost, and contingent valuation approaches to estimate the value of recreational fishing in the Swan River drainage compared with other sites in Montana. The data was obtained from on-site interviews and a creel census administered to users at sites along the Swan River, Swan Lake, and their tributaries. (*6 comparison ratios*)

Bojö (1985) undertook a study for Sweden's Environmental Protection Agency to estimate the benefits of protecting an area (called a Nature Reserve) from forest harvesting in the Vaalaa Valley in Northern Sweden. Bojö interviewed Vaalaa Valley visitors to gather information to estimate an average willingness to pay per visitor using both a travel cost and a contingent valuation model. (*1 comparison ratio*)

Devlin (1985) estimated the benefits from firewood collection in Northern Colorado National forests in his Ph.D. dissertation. A mail survey was used to collect the data. The author presented a travel cost estimate derived from an individual-observation model and two contingent valuation estimates, one based on willingness to pay and the other based on willingness to drive. (*2 comparison ratios*)

Donnelly, Loomis, Sorg and Nelson (1985) conducted a study in 1982 that estimated the average net willingness to pay for steelhead fishing trips in Idaho using both travel cost and contingent valuation methodologies. The data was obtained from a random sample of anglers purchasing Idaho steelhead fishing tags. The authors present per trip estimates derived from a multiple-site travel cost model and a contingent valuation model. (*1 comparison ratio*)

Michaelson and Smathers (1985) conducted a study to value camping and other outdoor recreation activities in the Sawtooth National Recreation Area using the travel cost and contingent valuation methodologies. The study data was obtained from on-site interviews of tourists and local users. The authors present a travel cost estimate and, using different payment vehicles, three contingent valuation estimates. (*3 comparison ratios*)

O'Neil (1985) estimated consumer surplus for recreation associated with two sites in Maine, the West Branch of the Penobscot River and an area of the Saco River. A contingent valuation and single-site travel cost analysis were conducted using information gathered from in-person interviews of site visitors conducted during the summer of 1984. Under different assumptions regarding functional form, the authors present travel cost estimates for each site and, using two different elicitation methods, contingent valuation estimates for each site. (*16 comparison ratios*)

Sellar, Stoll and Chavas (1985) used a regional travel cost model and two different contingent valuation elicitation approaches to estimate the value of recreational boating on four lakes in East Texas. The study data was collected using a questionnaire mailed to a sample of registered pleasure-boat owners in a 23-county area of East Texas. The authors present travel cost estimates for each of the four lakes, three contingent valuation estimates for three of the four lakes, and one contingent valuation estimate for the fourth lake. Net willingness to pay was calculated by subtracting average boat launch fees from estimates of gross consumer surplus. This procedure resulted in negative CV willingness to pay estimates in two instances. To avoid numeric complications, we set these CV/RP values equal to the smallest positive CV/RP ratio in the sample. (*10 comparison ratios*)

Walsh, Sanders and Loomis (1985) used a multi-site travel cost model and a contingent valuation model to value visits to a group of eleven rivers recommended for protection under the Wild and Scenic Rivers Act and to a second group of rivers. The data was obtained from a mail survey administered to a random sample of Colorado residents in 1983. (*4 comparison ratios*)

Wegge, Hanemann and Strand (1985) valued marine recreational fishing in Southern California using the contingent valuation methodology and multiple-site travel cost analysis. Their data was obtained from a 1984 mail survey of anglers. Separate estimates were derived for several modes of fishing: shore fishing, party/charter boat fishing, rental boat fishing and private boat fishing. The authors present both travel cost and contingent valuation estimates under different assumptions regarding functional form. (*42 comparison ratios*)

Loomis, Sorg and Donnelly (1986) estimated per-trip consumer surplus for cold-water fishing in Idaho. The authors used a multiple-site travel cost model and the contingent valuation method in estimation. The data was gathered in 1983 from a survey of anglers that elicited information about the anglers' fishing activity in 1982. Fifty-one sites were included in the study area. Travel cost estimates are provided using the full fifty-one sight model, a three site model, and a one-site model. A per-trip contingent valuation estimate is also provided. (*6 comparison ratios*)

Milon (1986) valued the construction of an artificial reef in Southern Florida. He estimated several different multi-site travel cost models that differed in the functional form used and the assumptions made about possible site-substitution. Using three subsamples with different elicitation methods, he obtained comparable contingent valuation data from a large mail survey. (*15 comparisons*).

Mitchell and Carson (1986) valued water pollution control using the contingent valuation methodology. The data was obtained from a large, national, in-person survey administered in 1983. The authors provide a comparison of their contingent valuation estimate with a travel cost estimate from Vaughan and Russell (1982). *(1 comparison ratio)*

Sorg and Nelson (1986) conducted a study to value elk hunting in Idaho using both travel cost and contingent valuation methodologies. A telephone survey was administered to resident and nonresident elk hunters holding a general elk hunting license in January and February of 1983 and 1984 to gather data on the 1982 and 1983 elk hunting seasons. Using standard and reported costs per mile, the authors present two travel cost estimates and two contingent valuation estimates. *(4 comparison ratios)*

Farber and Costanza (1987) conducted a study to estimate the social value of the Terrebonne Parish wetland system in South Louisiana. The social value of the wetland system was divided into three primary components: commercial fishing and trapping, recreation, and storm wind damage protection. The value of the wetlands area for one of these three components of value, recreation, was estimated using both travel cost and contingent valuation methodologies. The study data was obtained from a mail survey of Terrebonne wetland users that was administered July, 1984 through July, 1985. Under various assumptions regarding the value of travel time, three different travel cost estimates are compared to the contingent valuation estimate. *(3 comparison ratios)*

Hanley and Common (1987) used a zonal travel cost model and a contingent valuation model to estimate the recreational benefits derived by visitors to a part of the Queen Elizabeth Forest Park in Central Scotland. *(1 comparison ratio)*

Young *et al.* (1987) estimated the consumer surplus from small game hunting in Idaho. The authors provide an estimate for all upland game species and a separate estimate for pheasant hunting. Travel cost and willingness to pay information was gathered from a mail survey followed by a telephone survey of residents and nonresident licensed hunters. Travel cost estimates are reported using two different assumptions about the cost per mile. The authors provide four multiple-site travel cost model estimates and two contingent valuation estimates. *(4 comparison ratios)*

Adamowicz (1988) valued consumer surplus per day for hunting big horn sheep in six Alberta, Canada hunting zones. Twelve travel cost estimates and one contingent valuation estimate is provided for each hunting zone. The travel cost estimates are based on four different functional forms combined with three different time values. (*72 comparison ratios*)

Duffield and Allen (1988) used travel cost and contingent valuation methodologies to compare site-specific per-trip values for trout fishing on seventeen Montana rivers. The travel cost estimates are based on a 1985 survey while the contingent valuation estimates are based on the 1986 Angler Preference Survey administered by the Montana Department of Fish, Wildlife and Parks. The authors present site-specific estimates for each river derived from a multiple-site travel cost model and a contingent valuation model. (*17 comparison ratios*)

Navrud (1988, 1990, 1991a, 1991b) conducted four studies that all had a similar focus — estimating the recreational value per angling day using both travel cost and contingent valuation methodologies. Navrud's 1988 study valued freshwater fishing for salmon and sea trout at Norway's River Vikedalselv; the 1990 study valued, salmon and sea trout on the River Audna; and the 1991a study valued, brown trout at Lake Lauvann and Gjerstadskog Lakes (separate estimates are provided for each lake). Navrud's 1991b study valued saltwater fishing for salmon and sea trout at a sea area near River Audna. (*4, 4, 8, 4 comparison ratios*)

Ralston (1988) valued annual recreation benefits at Reelfoot Lake in Tennessee using the contingent valuation and travel cost methodologies. The data for both analyses was obtained in a mail-survey sent to lake visitors. A single-site, zonal travel cost model, with an internally generated value of time, provided one estimate. The contingent valuation estimate was based on an open-ended question for an annual pass to the lake. (*1 comparison ratio*)

Brown and Henry (1989) looked at the viewing of elephants on a wildlife safari tours in Kenya. They estimate both the contingent valuation a single-site travel cost model using several different assumptions. The CV data and some travel cost data was obtained from an on-site survey of tour participants. Most participants were from the United State or Europe. (*8 comparison ratios*)

Bockstael, McConnell, and Strand (1989) used the contingent valuation method and two variants of the travel cost method to estimate an improvement in the water quality of the Chesapeake Bay from current levels to an improved condition, one which the respondent considers acceptable for swimming. The data was obtained from a random sample of residents in the Baltimore-Washington SMSA. (*2 comparison ratios*)

Hanley (1989) conducted a study to value recreation benefits derived by visitors to a part of the Queen Elizabeth Forest park in Central Scotland using both travel cost and contingent valuation methodologies. The data was obtained in the summer of 1987 from in-person interviews and self-administered questionnaires of park visitors. Four travel cost estimates, based on different functional forms, are provided along with two contingent valuation estimates, one obtained from a close-ended format and another from an open-ended format. (*8 comparison ratios*)

Harley and Hanley (1989) conducted a study to value visits at the Loch Garten bird reserve in the Scottish Highlands using a semi-log travel cost model and a contingent valuation model. The data was obtained from on-site interviews of visitors to the reserve. (*1 comparison ratio*)

Huppert (1989) used travel cost and contingent valuation approaches to estimate the economic value associated with recreational fishing for chinook salmon and striped bass in Central California. The data used in the study was obtained from the Bay Area Sportfish Economic Survey which was carried out during 1985-86. Willingness to pay values were estimated using both the full sample and a subsample where respondents not catching any fish were dropped. (*4 comparison ratios*)

Johnson (1989) valued recreational fishing at two Colorado locations, Blue Mesa Reservoir and the Poudre River, using the travel cost and contingent valuation methodologies. Data for both analyses was obtained from a survey of visitors. For each location, two pairs of contingent valuation/travel cost estimates are provided. One pair is based on the maximum willingness to pay rather than forgo the recreational experience and the other is based on a change in catch. (*4 comparison ratios*)

Walsh, Ward and Olienik (1989) valued the effect of tree density on recreational demand for six recreational sites in Colorado. In the summer of 1980, in-person interviews were conducted to gather travel cost and contingent valuation data from site visitors. Contingent valuation estimates were derived using the full sample and a multiple-site travel cost model was estimated using a subsample of respondents. (*8 comparison ratios*)

White (1989) valued recreation at Belmar Beach in New Jersey using the contingent valuation and travel cost approaches. Data for both analyses was taken from a 1985 U.S. Army Corp of Engineers on-site survey of beach users. Two different single-site travel cost estimates are provided as well as three different CV estimates for four different sub-samples: season pass holders, day pass holders, season pass holders with summer residents excluded, and day pass holders with summer residents excluded. (*24 comparison ratios*)

Duffield and Neher (1990) estimated the consumer surplus associated with deer hunting in Montana using the contingent valuation method. The data used in their study was collected in a 1988 survey of hunters. The authors offer a comparison of their contingent valuation estimate with a travel cost estimate obtained from a companion study by Brooks (1988). (*1 comparison ratio*)

Richards, King, Daniel and Brown (1990) used travel cost and contingent valuation methodologies to estimate recreational consumer surplus for national forest campgrounds in northern Arizona. The data for the contingent valuation analysis was obtained from an on-site survey of recreationists at several national forest campgrounds in northern Arizona during the summer of 1985. The data for the travel cost analysis was compiled from fee envelopes collected by the U.S. Forest Service in 1985. The authors compare contingent valuation estimates with estimates derived from a multiple-site travel cost model for 10 campgrounds. (*10 comparison ratios*)

Walsh, Sanders and McKean (1990) used both travel cost and contingent valuation approaches to estimate a demand function for the recreation activity of pleasure driving/sightseeing by car along sections of eleven rivers in the Colorado Rocky Mountains. The data was obtained from a 1983 mail survey of Colorado's resident population. For per day consumer surplus, the authors present one travel cost and two contingent valuation estimates; for total trips, the authors present one travel cost and one contingent valuation estimate. (*3 comparison ratios*)

Willis and Garrod (1990) valued open-access recreation on inland waterways in the United Kingdom. Using data gathered from in-person interviews with canal users, the authors estimated recreational consumer surplus using the contingent valuation method and a multi-site travel cost model. A range of estimates is provided under different assumptions regarding functional form. (2 *comparison ratios*)

Loomis, Creel and Park (1991) valued deer hunting in California using the contingent valuation and travel cost methodologies. The data was obtained from a mail survey of California residents and nonresidents who had purchased a deer hunting license for the 1987 season. Under different assumptions regarding functional form, the authors present two travel cost estimates and compare those to a contingent valuation estimate. (2 *comparison ratios*)

Rolfsen (1991), estimated the recreational value per angling day for salmon and sea trout in Norway's freshwater Gaula River using the travel cost and contingent valuation methodologies. The author provides on contingent valuation estimate and a range spanned by travel cost estimates. We used the high and low estimates from this range. (2 *comparison ratios*)

Sievänen, Pouta and Ovaskainen (1991) valued recreation at a regional recreational area near Helsinki using a single-site travel cost model and two variants of the contingent valuation approach. The study data was obtained from on-site interviews of visitors to the recreational area. Willingness to pay and willingness to travel additional distances responses were elicited. (5 *comparison ratios*)

Duffield (1992) used data collected in an earlier study (Jones and Stokes, 1987) of sportfishing in Southcentral Alaska to estimate consumer surplus for sportfishing. The author estimated per trip consumer surplus using the contingent valuation method. Three contingent valuation estimates are provided: the estimated mean, an estimated truncated mean, and the estimated median. Also given is a multi-site travel cost mean estimate from Jones and Stokes (1987). Because a median travel cost estimate is not provided, we use only the estimated mean and estimated truncated mean for comparison with the travel cost estimate. (2 *comparison ratios*)

Mungatana and Navrud (1993) conducted a study to estimate the recreational value of wildlife viewing in Lake Nakuru National Park in Kenya using multi-site travel cost and contingent valuation approaches. The data was obtained from on-site interviews of park visitors in 1991. The authors present travel cost and contingent valuation estimates under different assumptions regarding functional form. Separate estimates were derived for flamingo viewing. (*6 comparison ratios*)

Choe, Whittington, and Lauria (1994) valued recreational benefits at an urban beach, which had been closed due to pollution, near Davao, Philippines using the contingent valuation and travel cost approaches. The data for both analyses was collected in an in-person survey in late 1992. One travel cost estimate is presented as well as four different CV estimates based on different forms of the valuation function. (*4 comparison ratios*)

3.2 Comparisons of Contingent Valuation with Hedonic Pricing

Darling (1973) valued amenities at three urban lakes in California using the hedonic pricing and contingent valuation approaches. The CV data was obtained from interviews of residents living in the areas surrounding the water parks. The hedonic price data was obtained from sales information and tax assessment records. For each of the three lakes, the author presented one hedonic price estimate and two CV estimates for the comparable categories. The CV estimates are derived using two different functional forms. (*6 comparison ratios*)

Loehman, Boldt and Chaikin (1981) valued changes in air quality in Los Angeles and the San Francisco Bay using contingent valuation and several different hedonic pricing approaches. The contingent valuation data was obtained from in-person interviews administered in areas with various air quality and socioeconomic characteristics for the two metropolitan areas in 1980. The authors provide one contingent valuation estimate for each metropolitan area; under different assumptions regarding functional form and pollution variables, they provide three hedonic price estimates for the Bay area and four hedonic price estimates for Los Angeles. (*7 comparison ratios*)

Brookshire, Thayer, Schulze and d'Arge (1982) used contingent valuation and hedonic pricing approaches to value improvements in air quality in Los Angeles. Property value information was gathered from a sample of single-family home sales that took place in late 1977 and early 1978. The contingent valuation data was obtained from in-person surveys administered in early 1978. (*11 comparison ratios*)

Gegax (1984) valued changes in risk using the contingent valuation and hedonic price methods. The data was obtained from a large, national mail-survey. The hedonic price estimate was obtained from the regression of log wages on respondent and occupational characteristics including several job-related risk variables. The contingent valuation estimate was obtained by using a payment card elicitation approach for a specified risk change. Gegax, Gerking and Schulze (1985) examined how much workers were willing to pay for job-related risk reduction using both hedonic pricing and contingent valuation methodologies. The authors used the same data as Gegax (1984) which was obtained from a national mail-survey during the summer of 1984. The authors present a range of value of life estimates from their contingent valuation results and a point estimate from their wage-risk analysis. (1,*2 comparison ratios*)

Blomquist (1984) explored the comparability of implicit market values by using a hedonic pricing mechanism and contingent market values to estimate two-view related amenities. Residents of ten high-rise buildings along the Lake Michigan shoreline in Chicago were interviewed in 1981 to obtain estimates for the value of both the lake view and high-rise view of a dwelling unit. Blomquist (1988) used a subset of this data to presented additional hedonic pricing and contingent valuation estimates. (*14, 4 comparison ratios*)

Brookshire, Thayer, Tschirhart and Schulze (1985) used contingent valuation and a hedonic pricing approach to study the value of an equivalent house inside and outside Los Angeles County's earthquake special study zones. The contingent valuation data was obtained from a survey of homeowners inside the special study zones that asked respondents about their willingness to pay for a potential transfer of home ownership from inside to outside the zone. The hedonic pricing data was based on a comparison completed before the passage of the Alquist-Prilo Act that designated the special study zones. (*1 comparison ratio*)

Pommerehne (1988) conducted a study to estimate willingness to pay for changes in road and aircraft noise using hedonic pricing and contingent valuation models. The data was obtained from in-person interviews administered to residents of Basle, Switzerland. The authors present estimates for changes in both road and aircraft noise. (*2 comparison ratios*)

d'Arge and Shogren (1989) conducted a study to value water quality in the Okoboji Lakes region of Iowa using CV and two variants of the hedonic pricing approach. The authors' contingent valuation study used data collected from a sample of area households in the summer and fall of 1984. One of the two hedonic price estimates is based on a standard hedonic price model and the other is based on a model derived from a survey of real estate professionals. (*3 comparison ratios)*

Randall and Kriesel (1990) valued 25 percent reductions in both air and water pollution in the United States. They used a discrete choice CV survey and a large multi-market hedonic pricing analysis. The authors present one CV estimate from a valuation function that pooled all data and included the survey mode as one of the explanatory variables. (*1 comparison ratio*)

3.3 Comparisons of Contingent Valuation with Averting Behavior/Household Production

Hill (1988) estimated the benefits of reducing the risk of breast cancer mortality in his PH.D. dissertation using contingent valuation and revealed preference methods. The survey data was obtained from a sample of women drawn from the Cancer Prevention Clinic, University of Wisconsin, Madison. The author presented revealed preference estimates of willingness to pay for an annual physical examination using two alternate model specifications and three risk groups and two contingent valuation estimates for comparable risk reductions. (*12 comparison ratios*)

Shechter, Kim and Golan (1989) valued the reduction of air pollution in the Haifa area of Northern Israel. Using information obtained from a 1986-1987 household survey, the authors estimated consumer surplus for the specified reductions using the contingent valuation method and an estimated demand system for health and housing services that are tied to air quality levels. Shechter (1991), using this dataset, breaks out contingent valuation estimates by three different elicitation methods and provides estimates from two variants of the household production approach, a health production and a consumer preference approach. Shechter and Kim (1991), again using the same data, present the same contingent valuation estimates as in

Shechter, Kim and Golan (1989) but now combined with some of the RP estimates from Shechter (1991). *(9, 2, 2 comparison ratios)*

John, Walsh, and Moore (1992) valued a Jefferson County, Texas mosquito abatement program using the contingent valuation method and an expenditure function approach. The data for both analyses was collected in a 1983 mail survey of Jefferson County residents. Benefits per household are provided using the two approaches. *(1 comparison ratio)*

3.4 Comparisons of Contingent Valuation with Actual/Simulated Markets

Bohm (1972) estimated the willingness to pay for the provision of a public television program in Sweden. Five groups were asked their willingness to pay with an explicit payment schedule provided to the respondents. Two other groups were asked to state their willingness to pay with no actual payment required. The study was conducted in November of 1969. *(10 comparison ratios)*

Kealy, Doridio, and Rockel (1987) conducted a study to estimate contingent values for preventing additional damages from acid rain to the Adirondack region's aquatic system. The authors' sample of undergraduates was divided into two treatments. In the first treatment, respondents were asked to make actual payments, and in the second, respondents were asked for their willingness to pay. The two treatments were administered in two sessions held two weeks apart. Actual estimates and CV scenario estimates were presented for each of the two sessions. *(2 comparison ratios)*

Hoehn and Fishelson (1988) valued consumer surplus associated with different visibility levels at the Hancock Tower Observatory in Chicago. Using actual attendance and visibility data, the authors used a conventional demand model and two quality adjusted demand models to value per trip consumer surplus for one particular visibility level. A contingent valuation survey was administered to Hancock Tower visitors to estimate consumer surplus for the same visibility improvement. Hoehn (1990) uses the same basic dataset and provides additional conventional demand model estimates and a well as a contingent valuation estimate for a different visibility level. *(3,3 comparison ratios)*

Sinden (1988) conducted four experiments concerning soil and forest conservation in Australia using the contingent valuation method and compared the results with actual contributions. The experiments were designed to test for potential information effects and hypothetical bias. Hypothetical willingness to pay for soil conservation and hypothetical willingness to pay for forest conservation were elicited. At the end of some of the experiments, respondents were given the opportunity to voluntarily donate into a fund marked for soil or forest conservation, thereby generating actual comparisons. *(17 comparison ratios)*

Boyce *et al.* (1989) conducted a study which is best known for its actual WTA and WTP experiments with Norfolk pine trees. However, one contingent valuation field study was conducted that elicited a respondent's willingness to pay for Norfolk pines under the condition that if the pine tree was not bought it would be killed. The authors compare this estimate to an estimate derived from a created market in which actual payments were required. *(1 comparison ratio)*

Bishop and Heberlein (1990) report on a 1983 and 1984 simulated-market, contingent-valuation field experiment in which respondents were able to purchase Wisconsin Sandhill Deer permits. The two 1983 experiments used an auction mechanism and the 1984 experiment used a discrete choice take-it-or-leave-it mechanism. The sample of individuals in each of the experiments were drawn from individuals who had expressed an interest in obtaining a permit and then randomly split into two subsamples. In the 1984 experiment one subsample was offered the opportunity to actually purchase a permit at a stated price while the other was asked, hypothetically, if they would purchase a permit at the stated price. *(3 comparison ratio)*

Duffield and Patterson (1991) conducted a field experiment in which one subsample of respondents of fishermen were asked for actual payment to help purchase water rights for Big and Swamp Creeks in Montana. Two other subsamples, which had different sponsors (*i.e.*, the University of Montana and the Montana Nature Conservancy), were asked about their willingness to pay, but no actual contribution was elicited. Comparisons were made for residents and nonresidents on a per contribution and per respondent basis. *(8 comparison ratios)*

Essenberg (1991) valued two different types of water systems in several different Philippine villages using the contingent valuation method. The contingent valuation data was obtained using in-person interviews that provided an iterative-bidding game elicitation format to respondents. One of the villages used in the contingent valuation survey was matched by characteristics with another village which had recently installed one of the described water systems. The author offers a comparison between the CV estimate and the actual payments made in the control village. (*1 comparison ratio*)

4. SUMMARIZING THE CV/RP RATIOS

Table 1 summarizes the CV/RP ratios treating the dataset in three different ways. The full sample uses each individual CV/RP ratio as an observation. The trimmed sample uses the remaining data after trimming off the smallest 5% and largest 5% of the CV/RP ratios. The weighted sample uses the mean CV/RP ratio from each study as that study's observation.[7] This weighting scheme prevents studies with multiple comparisons from having a disproportionate influence relative to studies reporting only one or a small number of comparisons. For each of the three treatments, we have provided the mean, the standard error of the mean, the maximum and minimum observations, the median (the 50th percentile), a wide range of other percentiles of the sample distribution, and finally, the sample size.

For the complete sample the estimate of the mean CV/RP ratio is 0.876 with a 95% confidence interval [0.800-0.952] and a median ratio of 0.739. For the trimmed sample, the estimate of the mean CV/RP ratio is 0.770 with a 95% confidence interval [0.729-0.811] and a median of 0.741.[8] For the

[7]The differences between the estimates from this treatment of the data and the complete and trimmed samples are due largely to the weighting (using the mean of each study's ratios) which reduces the influence of studies that provide multiple estimates. Adamowicz (1988) accounts for 72 comparisons; Desvousges, Smith and McGivney (1983) combined with estimates from Smith, Desvousges, and Fisher (1986) account for 48 comparisons (both use the same data); Wegge, Hanemann, and Strand (1985) account for 42 comparisons; White (1989) accounts for 24 comparisons; as well as eleven other studies that provide between 10 and 17 comparisons. Because we are considering ratios which are bounded below by zero and unbounded above, averaging is still understandably sensitive to large ratios within studies.

[8]Some of the most extreme variation in the CV/RP ratios come from a small number of studies and are subject to several qualifications: Smith, Desvousges, and Fisher (1986) (5 of the 10 largest ratios and 8 of the 10 smallest ratios) whose purpose was to pick assumptions which demonstrated how the analyst's judgement plays a very important role in the development of both CV and TC estimates; two of the Shechter papers (4 of the largest 10 ratios) used the same RP estimate, which was one-tenth and one-twentieth the size of the other two RP estimates for the same change, to compare with different CV estimates; Sellar, Stoll, and Chavas (1985) (2 of 10 smallest ratios) obtained negative net willingness to pay values; and ECO Northwest (1984) where the CV estimate was 5 times higher than one of the RP estimates, but one-half the size of the other RP estimate.

weighted sample the mean CV/RP ratio is 0.917 with a 95% confidence interval [0.813-1.021] and a median of 0.939.[9]

Figure 1 depicts a non-parametric density estimate of the full sample using a simple kernel density estimator first proposed by Wegman (1972; see also Silverman, 1986 and Statistical Sciences, Inc., 1993) with a width parameter of 0.5. Almost all of the density falls below a CV/RP value of 2.0 with almost 70% of the mass to the left of a CV/RP ratio of 1.0. This figure also shows a fairly long, but very shallow, right tail that would be even longer (to just past 10) if we had not cut it off at 6, which is the first time the density estimate has a relative frequency of zero. Figure 2 depicts the non-parametric density estimate for the trimmed sample. Because the maximum CV/RP ratio is slightly greater than 2.5, one can see that almost all of the density lies to the left, of 1.5 with over 80% to the left of 1.25. Figure 3 depicts the nonparametric density estimate for the weighted sample. This figure shows a very pronounced peak at about 1.0, with over half the density to the left and a thicker, but much shorter, right tail than Figure 1.

The analysis provided is not invariant to whether the CV estimate is chosen as the numerator of the ratio (as above) or as the denominator. One could instead have looked at the ratio of the RP to CV estimates. Here, for the complete dataset, one gets a mean value of 6.164 with a 95% confidence interval of [4.484-7.843] and a median estimate of 1.352. This estimate, which suggests that the RP estimates are on average over six times the CV estimates, is driven by the several large outliers noted earlier. Using the trimmed dataset, we estimate a smaller but still large mean RP/CV ratio of 2.835 with a 95% confidence interval [2.502-3.167]. For the weighted sample, the mean RP/CV ratio is 3.499 with a 95% confidence interval of [1.928-5.070] and a median of 1.335. Thus, looking at the RP/CV ratios suggests that RP estimates are on average considerably larger than their CV counterparts. Another way to say this is that if one drew a ratio at random from any of the three datasets summarized in Table 1, there is almost a 70%

[9]An alternative weighting scheme which is more robust to large outliers and also avoids giving disproportionate influence to studies with multiple estimates is to use the median ratio from each study (rather than the mean). Doing this results in a N=79 dataset of CV/RP ratios with mean 0.824, a 95% confidence interval [0.727-0.922] and a median of 0.858.

TABLE 1
CV/RP ESTIMATES FOR THREE SAMPLE TREATMENTS

Percentile	Complete Sample	Trimmed Sample	Weighted Sample
Mean	0.876	0.770	0.917
Standard Error	0.039	0.021	0.053
Maximum	10.269	2.071	2.506
99%	4.864	1.964	2.506
95%	2.071	1.655	1.893
90%	1.573	1.380	1.575
80%	1.204	1.151	1.153
75%	1.128	1.091	1.111
70%	1.048	1.016	1.096
60%	0.913	0.888	0.990
50%	0.739	0.741	0.939
40%	0.599	0.610	0.792
30%	0.456	0.485	0.640
25%	0.360	0.408	0.585
20%	0.239	0.332	0.568
10%	0.090	0.119	0.337
5%	0.042	0.088	0.170
1%	0.009	0.054	0.134
Minimum	0.005	0.043	0.134
N	541	486	79

chance of getting a CV/RP ratio less than 1.0, and there are some very small CV/RP ratios in the dataset which in turn imply very large RP/CV ratios.

We regressed the CV/RP ratios from the trimmed dataset on a set of dummy variables representing the RP technique used with the single site travel cost models (TC1) as the omitted category. These results are shown in Table 2 with the t-statistics reported based on the White (1980) heteroskedasticity-consistent covariance matrix. They suggest the CV estimates run about 20% lower than the TC1 counterparts, about 30% lower than their TC2 counterparts, a little less than 40% lower than their HP counterparts, about 20% lower than their AVERT counterparts, and are, on average, indistinguishable from their ACTUAL counterparts. We also regressed the CV/RP ratios from the trimmed dataset on a set of dummy variables for the broad class of goods valued. These results are shown in Table 3 with the t-statistics similarly calculated. They suggest that the HEALTH goods may have CV/RP ratios closer to 1.0 relative to the other two categories of goods, although this conclusion should be tempered by the smaller number of CV/RP estimates in the HEALTH category[10].

An obvious next step is a more detailed analysis of this data using additional variables which show the specific details of the contingent valuation implementation, and a finer partitioning of the RP techniques and indicators of reliability such as sample size and standard errors. However, our efforts to conduct this analysis have been greatly hindered by the curse suffered by other meta-analyses of non-market data (*e.g.*, Smith and Karou, 1990): incomplete reporting of the necessary details. With rapidly declining sample sizes due to missing data and a large set of dummy variables, we found we were essentially identifying individual studies with particularly large or small CV/RP ratios. However, some general observations may be warranted. Many of these are along the lines of the meta-analyses of contingent valuation, travel cost analysis, and hedonic pricing which have been performed previously (Smith and Karou, 1990; Walsh, McKean, and Johnson, 1992; Smith and Huang, 1993; Smith and Osborne, 1994). The single-site travel

[10]Results based on the full data set are quite similar in both relative and absolute magnitude for the various RP techniques with the exception of TC1, represented by the intercept term, which is 0.9719, and AVERT which has a significant positive coefficient. Neither the HEALTH nor ENVAM dummies are significant in the regression equation using the complete dataset.

TABLE 2
REGRESSION OF CV/RP on RP TECHNIQUE USED

Parameter	Estimate	t-Statistic
vIntercept	0.8030	24.38
TC2	-0.1206	-2.38
HP	-0.1876	-2.84
AVERT	0.0064	0.09
ACTUAL	0.2332	3.73
N=486	R^2=.06	

TABLE 3
REGRESSION OF CV/RP on TYPE OF GOOD VALUED

Parameter	Estimate	t-Statistic
Intercept	0.7629	30.01
ENVAM	0.0003	0.01
HEALTH	0.2223	1.97
N=486	R^2=.01	

cost models produce higher CV/RP ratios on average than do the multiple-site models. This is largely because many of the TC1 models do not include any value of travel time while most TC2 models make some allowance for travel time cost. The TC2 models also tend to be more elaborate with some visitors coming from long distances to one or more of the sites examined. Estimates from the TC2 are often presented using different functional forms, some of which produce quite large RP numbers. Hedonic pricing and averting/household production models are quite sensitive to the particular functional form and attributes used, and can generate a wide range of RP estimates from the same dataset. The CV estimates vary with the treatment of outliers and protest responses, the functional form used with discrete choice CV data, and the payment mechanism used. CV estimates are undoubtedly sensitive to how well the good is described and whether the respondents believe the good can be provided (Mitchell and Carson, 1989). RP estimates are undoubtedly sensitive to the researcher's assumptions about a good's input costs (Randall, 1994) and characteristics (Freeman, 1993).[11]

5. CORRELATION BETWEEN CV AND RP ESTIMATES

The average CV/RP ratio does not directly address whether CV and RP estimates tend to move together.[12] The convergent validity of the two measurement techniques is closely tied to the presence of a significant correlation between the estimates derived using the different techniques, although one can argue about how large such a correlation needs to be. For the complete sample, the correlation coefficient is 0.40. For the trimmed sample, the correlation coefficient is higher, 0.60, as would be expected since we have trimmed out the small number of highly divergent CV/RP estimates. For the weighted data set, the correlation coefficient is even larger, 0.68, which also makes sense since this sample is obtained by averaging multiple CV/RP ratios which often fall on both sides of one. In all three instances, the correlation coefficient is significantly different ($p < 0.001$) from zero.

---

[11]For instance, recreationists' costs of travel may differ greatly from the researcher's assigned costs or lake users may be unaware of an invisible toxin known to the researcher. In both cases, there is a divergence between the researchers's assumptions and the consumer's perceptions.

[12]It is possible to have an average CV/RP ratio of 0.5 or 2.0 and have the correlation between the two estimates be 1.0. Although it is unlikely with a large sample size, it is also possible to have an average CV/RP ratio of 1.0 and a correlation coefficient of zero.

6. OTHER COMPARISON APPROACHES

Comparing willingness to pay estimates from contingent valuation to estimates from revealed preference methodologies is certainly the most popular way of comparing the two approaches, but it is not the only one. One simple approach is to compare estimates of the fraction of a particular population who say that they will undertake a given activity with the fraction who actually undertake the activity. For example, Carson, Hanemann, and Mitchell (1987) look at the correspondence between the estimate of the percent who say in a survey that they will vote for a water quality bond issue (70-75%) and the percent actually voting in favor of it (73%). Kealy, Montgomery and Dovidio (1990) find that 72% of those who said they would donate money to the New York Department of Environmental Conservation to reduce acid rain in the Adirondacks actually did so several weeks later. This percentage increased to 92% in a subsample in which they strongly stressed the future payment obligation.[13] In contrast, Seip and Strand (1991), using members of a Norwegian environmental group as interviewers, find that only 10% of respondents who indicate they would be willing to pay a specified membership fee for the group actually did so when solicited a month later. Navrud (1992) conducted a similar exercise, but this time sampling people who had sent in a reply coupon from a full page World Wildlife Federation (WWF) newspaper ad in Norway "contributing their vote as a WWF friend." While Navrud's study showed several times the percentage joining the environmental group as Seip and Strand's study, what Navrud emphasizes is the difficulty of drawing a close correspondence between a vague initial request which potentially confuses support for the environmental group's public goals with the actual private good purchase of membership in the group.

Analysts are also often interested in other economic quantities such as elasticities. For example, Cummings *et al.* (1986) estimate the elasticity of substitution between wages and municipal infrastructures in western boomtowns to be -0.35 using a hedonic wage equation estimated on data from 29 towns and -0.037 to -0.042 using CV surveys done in three boomtowns. Thomas and Styme (1988) used a contingent

[13]The number of subjects who declined to donate after earlier saying they would was only slightly larger than the number of subjects who said they would not donate but who actually did so.

valuation study in Perth, Australia to estimate the residential water demand price elasticity since there had been little prior variation in water rates. They estimated the price elasticity to be -0.20, whereas econometric models estimated after changes in the water rates put the price elasticity in the range of -0.10 to -0.43. In the public finance literature, income and tax price elasticities for a particular good estimated from survey choice data have compared favorably to those estimated from aggregate voting data and governmental provision decisions (Bergstrom, Rubinfeld, and Shapiro, 1982; Gramlich and Rubinfeld, 1982).

A different approach is to compare the utility of different choices from stated preference (SP) and RP models using the suggestions of Louviere and Timmermans (1990) for recreational modeling.[14] In some instances, it may also be possible to compare parameters estimated from different models (Hensher *et al.* 1989). Hensher uses this approach to show the similarity of the value of travel time estimates from the two types of models in the transportation literature. With adequate and similar information on the variables underlying the choice process, one can directly test for the statistical equivalence of the estimated contingent valuation and revealed preference choice models. Mu (1988) shows this for the choice problem of where to obtain household water in Brazil.[15] A less structured approach that is based on the non-parametric consumer preference framework of Varian (1983) has been applied to contingent valuation and travel cost data for big horn sheep hunting in Canada by Adamowicz and Graham-Tomasi (1991) to examine the consistency of the data underlying each of the two approaches with the basic theoretical restrictions on consumer demand. They show that most of their data from both approaches is consistent with the basic set of theoretical restrictions, with the contingent valuation data showing fewer violations.

---

[14]Utilities from choice models estimated from RP and SP data cannot be directly compared unless one takes account of the possibility of different latent scale parameters underlying the choice models (Morikawa, 1989). A number of comparisons in the literature which were previously thought to be divergent have been shown to be consistent once differences in scale (which is related to reliability) are taken into account.

[15]It is difficult to test whether the CV and RP data were generated by the same utility function without making strong structural assumptions about the choice process. It is particularly difficult unless one has obtained they key variables underlying that process for both the RP and CV samples. See Larson (1990) for an application and discussion of problems with this approach.

If one is prepared to say that neither CV nor RP data is inherently superior to the other, an obvious thing to do is combine them in some fashion. This approach has seen several recent applications in the marketing and transportation literatures (Ben-Akiva and Morikawa, 1990; Hensher and Bradley, 1993; Swait and Louviere, 1993), and has seen some initial applications (Adamowicz, Louviere, and Williams, forthcoming; Cameron, 1992; Hanemann, Chapman and Kanninen, 1993) in the recreational demand literature. Cameron (1992) has proposed an interesting scheme for looking at the results of differentially weighting the two sources of information in terms of consumer surplus estimates.

7. CONCLUDING REMARKS

Our examination of over 79 studies containing 541 CV/RP comparisons for quasi-public goods finds that CV estimates are smaller, but not grossly smaller, than their RP counterparts. For the complete dataset, 1.0 is just outside the upper-end of the 95% confidence interval [0.80-0.95] for the mean CV/RP ratio (0.88).[16] For the trimmed dataset, one can reject the hypothesis that the mean CV/RP ratio (0.77) is 1.0 in favor of the alternate hypothesis that it is less than one. For the weighted dataset, the mean CV/RP ratio (.92) is not significantly different from 1.0 using a 5% two-sided t-test. The median CV/RP ratios range between 0.74 and 0.94 depending upon the treatment of the sample. Most of the density lies in the range of CV/RP ratios of 0.25 to 1.25. The correlation coefficient between the CV and RP estimates varies between .40 and .68, depending on the sample considered, and is always significant at $p > .001$, thus providing some support for the convergent validity of the two basic approaches to non-market valuation of quasi-public goods.

Should these results be taken to imply that, CV estimates on average are equal to or slightly less than their RP counterparts? Our preliminary results suggest so. However, some CV estimates clearly exceed their revealed preference counter-parts, and therefore one cannot conclude that CV estimates are

[16]By carefully selecting a smaller number of studies, one could argue either that the CV/RP ratio was almost always 1.0 or that it was almost always substantially larger or smaller than 1.0. While one may want to choose a subset of the studies considered here on the basis of the quality of the study or to choose a restricted subset of the CV/RP comparisons from a particular study, it is important to carefully justify such choices.

always smaller than revealed preference estimates. Nonetheless, based on the available CV/RP comparisons, discounting CV estimates by a factor of two or more, as some have proposed, appears to be unwarranted given that CV/RP ratios of greater than 2.0 comprise only 5% of our complete sample and only 3% of our weighted sample. Indeed, applying a discount factor of 2.0 or greater to the CV estimates used in our analysis would result in "adjusted" CV estimates that, in almost all cases, *diverge* from the estimates obtained from observable behavior, rather than *converge*.

References

1. Adamowicz, W.L. "Extramarket Benefit, Evaluation and Revealed Preference." Dissertation, *DAI* 50, 1988, 495. University of Minnesota.

2. Adamowicz, W.L., J.J. Louviere and M. Williams. "Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities." *Journal of Environmental Economics and Management*, forthcoming.

3. Adamowicz, W.L. and T. Graham-Tomasi. "Revealed Preference Tests of Nonmarket Goods Valuation Methods." *Journal of Environmental Economics and Management* 20, 1991, 29-45.

4. Beardsley, W. *Economic Value of Recreation Benefits Determined by Three Methods*, Research Note RM-176. Rocky Mountain Forest and Range Experiment Station, Colorado Springs: U.S. Forest Service, 1970.

5. Ben-Akiva, M. and T. Morikawa. "Estimation of Switching Models from Revealed Preferences and Stated Intentions." *Transportation Research* 24A, 1990, 485-495.

6. Berg, C.B. "Publication Bias," in *The Handbook of Research Synthesis*, edited by Harris Cooper and Larry V. Hedges. New York: Russell Sage Foundation, 1994.

7. Bergstrom, T.C., D.L. Rubinfeld and P. Shapiro. "Micro-Based Estimates of Demand Functions for Local School Expenditures." *Econometrics* 50, 1982, 1183-1205.

8. Binkley, C.S. and W.M. Hanemann. *The Recreation Benefits of Water Quality Improvement: Analysis of Day Trips in an Urban Setting*, Report to the U.S. Environmental Protection Agency. Washington, DC: U.S. Environmental Protection Agency, 1978.

9. Bishop, R.C. and T.A. Heberlein. "Measuring Values of Extra-Market Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics* 61, 1979, 926-930.

10. Bishop, R.C. and T.A. Heberlein. "The Contingent Valuation Method," in *Economic Valuation of Natural Resources: Issues, Theory, and Applications*, edited by Rebecca L. Johnson and Gary V. Johnson. Boulder: Westview Press, 1990.

11. Blomquist, G.C. "Measuring the Benefits of Public Goods Provision Using Implicit and Contingent Markets." Unpublished paper. College of Business and Economics, University of Kentucky, August, 1984.

12. Blomquist, G.C. "Valuing Urban Lakeview Amenities Using Implicit and Contingent Markets." *Urban Studies* 25, 1988, 333-340.

13. Bockstael, N.E., K.E. McConnell and I.E. Strand. "Measuring the Benefits of Improvement in Water Quality: The Chesapeake Bay." *Marine Resource Economics* 6, 1989, 1-18.

14. Bohm, P. "Estimating Demand for Public Goods: An Experiment." *European Economic Review* 3, 1972, 111-130.

15. Bohm, P. "Revealing Demand for an Actual Public Good." *Journal of Public Economics* 24, 1984, 135-151.

16. Bojö, J. *Kostnadsnyttoanalys av Fjällnära Skogar. Fallet Vålådalen (A Cost-Benefit Analysis of Mountainous Forests. The Case of the Vålå Valley).* Stockholm, Sweden: Stockholm School of Economics, 1985.

17. Boyce, R. *et al.* "Experimental Evidence of Existence Value in Payment and Compensation Contexts." Paper presented at the USDA W-133 Annual Meeting. San Diego, California, February, 1989.

18. Brooks, R.. *The Net Economic Value of Deer Hunting in Montana*, Report prepared for Montana Department of Fish, Wildlife, and Parks. Helena, MT: State of Montana. Department of Fish, Wildlife, and Parks, 1988.

19. Brookshire, D.S., M.A. Thayer, J. Tschirhart and W.D. Schulze. "A Test of the Expected Utility Model: Evidence from Earthquake Risks." *Journal of Political Economy* 93, 1985, 369-389.

20. Brookshire, D.S., M.A. Thayer, W.D. Schulze and R.C. d'Arge. "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches." *American Economic Review* 72, 1982, 165-177.

21. Cameron, T.A.. "Combining Contingent Valuation and Travel Cost Data for the Valuation of Nonmarket Goods." *Land Economics* 68, 1992, 302-317.

22. Carson, R.T., W.M. Hanemann and R.C. Mitchell. "The Use of Simulated Political Markets to Value Public Goods." Discussion paper 87-7. Department of Economics, University of California-San Diego, 1987.

23. Carson, R.T., J.L. Wright, N.J. Carson, A. Alberini, N.E. Flores. *A Bibliography of Contingent Valuation Studies and Papers*, Bibliography submitted to the NOAA Blue Ribbon Panel on Contingent Valuation. Department of Economics, University of California, San Diego, 1994.

24. Chapman, D., B. Kanninen and W.M. Hanemann. "Non-Market Valuation Using Contingent Behavior: Model Specification and Consistency Tests." Paper presented at the Association of Environmental and Resource Economist Meeting. Anaheim, CA, January, 1993.

25. Choe, K., D. Whittington, and L.T. Donald. "The Economic Benefits of Surface Water Quality Improvements in Developing Countries: A Case Study of Davao, Philippines." Draft report to the Environment Department, the World Bank. University of North Carolina, Chapel Hill, March, 1994.

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

27. Cummings, R.G., W.D. Schulze, S. Gerking and D.S. Brookshire. "Measuring the Elasticity of Substitution of Wages for Municipal Infrastructure: A Comparison of the Survey and Wage Hedonic Approaches." *Journal of Environmental Economics and Management* 13, 1986, 269-276.

28. Darling, A.H. "Measuring Benefits Generated by Urban Water Parks." *Land Economics* 49, 1973, 22-34.

29. d'Arge, R.C. and J.F. Shogren. "Non-Market Asset Prices: A Comparison of Three Valuation Approaches," in *Valuation Methods and Policy Making in Environmental Economics*, edited by H. Folmer and E. van Ireland. Amsterdam: Elsevier Science Publishers, 1989.

30. Desvousges, W.H., V.K. Smith and M.P. McGivney. *A Comparison of Alternative Approaches for Estimating Recreation and Related Benefits of Water Quality Improvements*, EPA Report 230-05-83-001. U.S. Environmental Protection Agency. Office of Policy Analysis, Washington, DC, 1983.

31. Devlin, P.J. "Quantifying Recreation: Allocation of Value to the Collection and Burning of Firewood in Colorado." Dissertation, *DAI* 46, 1985, 1735. Colorado State University.

32. Diamond, P.A. and J.A. Hausman. "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives*, Forthcoming.

33. Donnelly, D.M., J.B. Loomis, C.F. Sorg and L.J. Nelson. *Net Economic Value of Recreational Steelhead Fishing in Idaho*, U.S. Forest Service Resource Bulletin RM-9. Fort Collins, CO: U.S. Forest Service. Rocky Mountain Forest and Range Experiment Station., 1985.

34. Duffield, J.W. "Travel Cost and Contingent Valuation: A Comparative Analysis," in *Advances in Applied Microeconomics*, volume 3, edited by V. Kerry Smith and Ann D. Witte. Greenwich, CT: JAI Press, 1984.

35. Duffield, J.W. "Exxon Valdez-Lost Recreation Use," in *Natural Resource Damages: Law and Economics*. New York: Wiley Law Publications, 1992.

36. Duffield, J.W. and S. Allen. "Contingent Valuation of Montana Trout Fishing by River and Angler Subgroup." Final Economic Report prepared for Montana Department of Fish, Wildlife, and Parks. Helena, MT: Montana Department of Fish, Wildlife, and Parks, September, 1988.

37. Duffield, J.W. and C. Neher. *A Contingent Valuation Assessment of Montana Deer Hunting: Attitudes and Economic Benefits*, Report to Montana Department of Fish, Wildlife, and Parks. Helena, MT: Montana Department of Fish, Wildlife, and Parks, October, 1990.

38. Duffield, J.W. and D.A. Patterson. "Field Testing Existence Values: An Instream Flow Trust Fund for Montana Rivers." Paper presented at the annual meeting of the Association of Environmental and Resource Economics. New Orleans, January, 1991.

39. ECO Northwest. *Economic Valuation of Potential Losses of Fish Populations in the Swan River Drainage*, Final report prepared for the Montana Department of Fish, Wildlife, and Parks. Helena, MT: Montana Department of Fish, Wildlife and Parks, September, 1984.

40. Essenburg, T.J. "Financing Publicly Provided Goods in Less-Developed Countries: A Contingent Valuation Study of Potable Water Services in Rural Philippines." Dissertation, *DAI* 52, 1991, 4412. University of Tennessee.

41. Farber, S.C. and R. Costanza. "The Economic Value of Wetlands Systems." *Journal of Environmental Management* 24, 1987, 41-51.

42. Freeman, A.M., III. *The Measurement of Environmental and Resource Values*. Washington, DC: Resources for the Future, 1993.

43. Gegax, D.A. "Alternative Measures of the Value of Safety: A Theoretical and Empirical Comparison." Dissertation, *DAI* 46, 1984, 1038. Department of Economics, University of Wyoming.

44. Gegax, D.A., S. Gerking and W.D. Schulze. "Perceived Risk and the Marginal Value of Safety." *Review of Economics and Statistics* 73, 1991, 589-596.

45. Gramlich, E. and D. Rubinfeld. "Micro Estimates of Public Spending Demand Functions and Tests of the Tiebout and Median Voter Hypotheses." *Journal of Political Economy* 90, 1982, 536-560.

46. Hanley, N.D. "Valuing Rural Recreation Benefits: An Empirical Comparison of Two Approaches." *Journal of Agricultural Economics* 40, 1989, 361-374.

47. Hanley, N.D. and M.S. Common. *Estimating Recreation, Wildlife and Landscape Benefits Attached to Queen Elizabeth Forest Park*, Final report to Forestry Commission. Edinburgh, Great Britain, 1987.

48. Harley, D. and Nicholas D. Hanley. "Economic Benefit Estimates for Nature Reserves: Methods and Results." Discussion Paper 89-6. Department of Economics, University of Stirling, Great Britain, 1989.

49. Harris, C.C. "Assessing the Validity of Economic Methods for Evaluating Sport Fishery Benefits: A Behavioral Approach." Dissertation, *DAI* 44, 1983, 577. University of Michigan.

50. Haspel, A.E. and F.R. Johnson. "Multiple Destination Trip Bias in Recreation Benefit Estimation." *Land Economics* 58, 1982, 364-372.

51. Hensher, D.A., P. Barnard, F. Milthorpe and N. Smith. "Urban Tollways and the Valuation of Travel Time." *Economic Record* 66, 1989, 146-156.

52. Hensher, D.A. and M. Bradley. "Using Stated Response Choice Data to Enrich Revealed Preference Discrete Choice Models." *Marketing Letters* 4, 1993, 139-151.

53. Hill, J.W. "Valuing Reductions in the Risk of Breast Cancer Mortality: A Comparison of Revealed Preference and Contingent Valuation Methods." Dissertation, *DAI* 50, 1988, 2992. University of Wisconsin, Madison.

54. Hoehn, J.P. "Contingent Valuation and the Prospect of a Satisfactory Benefit-Cost Indicator," in *Economic Valuation of Natural Resources*, edited by Rebecca L. Johnson and Gary V. Johnson. Boulder: Westview, 1990.

55. Hoehn, J.P. and G. Fishelson. "Visibility and Urban Recreation," in *The Economic Value of Visibility*, edited by G.S. Tolley and R. Fabian. Mt. Pleasant, MI: The Blackstone Co., 1988.

56. Hoehn, J.P. and A. Randall. "A Satisfactory Benefit Cost Indicator from Contingent Valuation." *Journal of Environmental Economics and Management* 14, 1987, 226-247.

57. Huppert, D.D. "Measuring the Value of Fish to Anglers: Application to Central California Anadromous Species." *Marine Resource Economics* 6, 1989, 89-107.

58. John, K.H., R.G. Walsh and C.G. Moore. "Comparison of Alternative Nonmarket Valuation Methods for an Economic Assessment of a Public Program." *Ecological Economics* 5, 1992, 179-196.

59. Johnson, D.M. "Economic Benefits of Alternative Fishery Management Programs." Dissertation, *DAI* 51, 1989, 582. Colorado State University.

60. Johnson, D.M. and R.G. Walsh. *Economic Benefits and Costs of the Fish Stocking Program at Blue Mesa Reservoir, Colorado*, Technical Report No. 49. Fort Collins, CO: Colorado Water Resources Research Institute, Colorado State University, 1987.

61. Johnson, F.R. and A.E. Haspel. "Economic Valuation of Potential Scenic Degradation at Bryce Canyon National Park." U.S Naval Academy and the Office of Policy Analysis, U.S. Department of the Interior, 1983.

62. Jones and Stokes Associates. *Southcentral Alaska Area Sport Fishing Economic Study*, Final report to the Alaska Department of Fish and Game. Sacramento: Jones and Stokes Associates, 1987.

63. Kealy, M.J., J.F. Dovidio and M.L. Rockel. "On Assessing the Magnitude of Bias in Contingent Values." Unpublished paper. Colgate University, Hamilton, NY, August, 1986.

64. Kealy, M.J., M. Montgomery and J.F. Dovidio. "Reliability and Predictive Validity of Contingent Values: Does the Nature of the Good Matter?" *Journal of Environmental Economics and Management* 19, 1990, 244-263.

65. Knetsch, J.L. and R.K. Davis. "Comparisons of Methods for Resource Evaluation." Water Research, edited by A.V. Kneese and S.C. Smith. Baltimore, MD: Johns Hopkins University Press for Resources for the Future, 1966.

66. Larson, D.M. "Testing Consistency of Direct and Indirect Methods for Valuing Nonmarket Goods." Unpublished paper. University of California, Davis, December, 1990.

67. Loehman, E.T., D. Boldt and K. Chaikin. *Measuring the Benefits of Air Quality Improvements in the San Francisco Bay Area*, Final report. Menlo Park, CA: SRI International, May, 1981.

68. Loomis, J.B., M. Creel and T. Park. "Comparing Benefit Estimates from Travel Cost and Contingent Valuation Using Confidence Intervals for Hicksian Welfare Measures." *Applied Economics* 23, 1991, 1725-1731.

69. Loomis, J.B., C.F. Sorg and D.M. Donnelly. "Evaluating Regional Demand Models for Estimating Recreation Use and Economic Benefits: A Case Study." *Water Resources Research* 22, 1986, 431-438.

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

71. Michalson, E.L. and R.L. Smathers, Jr. "Comparative Estimates of Outdoor Recreation Benefits in the Sawtooth National Recreation Area, Idaho." Agricultural Economics Research Paper Series 249. Department of Agricultural Economics, University of Idaho, May, 1985.

72. Milon, J.W. "Strategic Incentives Revisited: Valuation Context and Specification Comparability for Local Public Goods." Paper presented at the annual meeting of the Association of Environmental and Resource Economics. New Orleans, December, 1986.

73. Mitchell, R.C. and R.T. Carson. *The Use of Contingent Valuation Data for Benefit-Cost Analysis in Water Pollution Control*, Report to the U.S. Environmental Protection Agency. Washington, DC: U.S. Environmental Protection Agency, September, 1986.

74. Mitchell, R.C. and R.T. Carson. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future, 1989.

75. Morikawa, T. "Incorporating Stated Preference Data in Travel Demand Analysis." Dissertation. Department of Civil Engineering, Massachusetts Institute of Technology, June, 1989.

76. Mungatana, E.D. and S. Navrud. "Environmental Valuation in Developing Countries: The Recreational Value of Wildlife Viewing." Unpublished paper. Department of Agricultural and Resource Economics, University of California, Berkeley, October, 1993.

77. Navrud, S. *Nytte-Kostnadsanalyse av Vassdragskalking: En Studie i Audna (Cost Benefit Analysis of River Liming: A Case Study of River Audna)*, Report 1990-6, in Norwegian. Norway: Directorate for Nature Management, 1990.

78. Navrud, S. *Rekreasjonsverdien av Lakse - og Sjoearuefisket i Vikedalselva før Regelmessig Kalking (Recreational Value of Atlantic Salmon and Sea-Trout Angling in the River Vikedalselv Before Regular Liming)*, Report from the Directorate of Nature Management. In Norwegian, summary in English. Oslo: Directorate for Nature Management, 1988.

79. Navrud, S. *Samfunnsøkonomisk lønnsomhet av å kalke utvalgte aurevann i Agderfylkene (Social Profitability of Liming Selected Trout Lakes in the Agder Counties)*, Report from the Directorate for Nature Management, in Norwegian. Directorate for Nature Management, Trondheim, Norway, 1991.

80. Navrud, S. *Samfunnsøkonomisk lønnsomhet av å kalke Audna. (Social Profitability of Liming River Audna. An Extended Analysis)*, Report to the Directorate for Nature Management, in Norwegian. Directorate for Nature Management, Trondheim, Norway, 1991.

81. Navrud, S. "Willingness to Pay for Preservation of Species - An Experiment with Actual Payments," in *Pricing the European Environment*, edited by Ståle Navrud. Oxford: Oxford University Press, 1992.

82. O'Neil, W.B. "Estimating the Recreational Value of Maine Rivers: An Experiment with the Contingent Value Technique." Unpublished paper. Department of Economics, Colby College, Waterville, ME, August, 1985.

83. Pommerehne, W.W. "Measuring Environmental Benefits: A Comparison of Hedonic Technique and Contingent Valuation," in *Welfare and Efficiency in Public Economics*, edited by Dieter Bös, Manfred Rose and Christian Seidl. New York: Springer, 1988.

84. Ralston, S.N. "The Direct and Indirect Estimation of Recreational Benefits for Reelfoot Lake." Dissertation, *DAI* 50, 1988, 753. University of Tennessee.

85. Randall, A. "A Difficulty with the Travel Cost Method." *Land Economics* 70, 1994, 88-96.

86. Randall, A. and W. Kriesel. "Evaluating National Policy Proposals by Contingent Valuation," in *Economic Valuation of Natural Resources: Issues, Theory, and Applications*, edited by Rebecca L. Johnson and Gary V. Johnson. Boulder: Westview Press, 1990.

87. Richards, M.T., D.A. King, T.C. Daniel and T.C. Brown. "The Lack of an Expected Relationship Between Travel Cost and Contingent Value Estimates of Forest Recreation Value." *Leisure Sciences* 12, 1990, 303-319.

88. Rolfsen, J. *Rekreasjonsverdien av Lakse - Og Sjoeaurefiske på Tofa's soner i Gaula Sesongen (Recreation Value of Atlantic Salmon and Sea Trout Angling in Parts of River Gaula*, M.S. Thesis. Department of Economics, University of Oslo, Norway, 1990.

89. Seip, K. and J. Strand. "Willingness to Pay for Environmental Goods in Norway: A Contingent Valuation Study with Real Payment." *Environmental and Resource Economics* 2, 1992, 91-106.

90. Sellar, C., J.R. Stoll and J.P. Chavas. "Validation of Empirical Measures of Welfare Changes: A Comparison of Nonmarket Techniques." *Land Economics* 61, 1985, 156-175.

91. Shechter, M. "A Comparative Study of Environmental Amenity Evaluations." *Environmental and Resource Economics* 1, 1991, 129-155.

92. Shechter, M. and M. Kim. "Valuation of Pollution Abatement Benefits: Direct and Indirect Measurement." *Journal of Urban Economics* 30, 1991, 133-151.

93. Shechter, M., M. Kim and L. Golan. "Valuing a Public Good: Direct and Indirect Valuation Approaches to the Measurement of the Benefits from Air Pollution Abatement," in *Valuation Methods and Policy Making in Environmental Economics*, edited by H. Folmer and E. van Ireland. Amsterdam: North-Holland, 1989.

94. Sievänen, T., E. Pouta and V. Ovaskainen. "Problems of Measuring Recreation Value Given Everyman's Rights." *Scandinavian Forest Economics* 32, 1991.

95. Silverman, B.W. "Density Estimation for Statistics and Data Analysis." London: Chapman and Hall, 1986.

96. Sinden, J.A. "Empirical Tests of Hypothetical Biases in Consumers' Surplus Surveys." *Australian Journal of Agricultural Economics* 32, 1988, 98-112.

97. Smith, N.E. *A Comparison of the Travel Cost and Contingent Valuation Methods of Recreation Valuation at Cullaby Lake County Park*, Thesis. Oregon State University, 1980.

98. Smith, V.K., W.H. Desvousges and A. Fisher. "A Comparison of Direct and Indirect Methods for Estimating Environmental Benefits." *American Journal of Agricultural Economics* 68, 1986, 280-290.

99. Smith, V.K. and J.C. Huang. "Hedonic Models and Air Pollution: Twenty-Five Years and Counting." *Environmental and Resource Economics* 3, 1993, 381-394.

100. Smith, V.K. and Y. Kaoru. "Signals or Noise? Explaining the Variation in Recreation Benefit Estimates." *American Journal of Agriculture Economics* 72, 1990, 419-433.

101. Smith, V.K. and L. Osborne. "Do Contingent Valuation Estimates Inform Environmental Policy?" Paper presented at the American Economic Association Meeting. Boston, January, 1994.

102. Sorg, C.F. and L.J. Nelson. *Net Economic Value of Elk Hunting in Idaho*, U.S. Forest Service Resource Bulletin RM-12. Fort Collins, CO: U.S. Forest Service. Rocky Mountain Forest and Range Experiment Station, 1986.

103. Statistical Sciences, Inc. *S-PLUS for Windows: Reference Manual.* Seattle: Statistical Sciences, Inc., 1993.

104. Swait J., Louviere J.J. "The Role of the Scale Parameter in the Estimation and Use of Generalized Extreme Value Models." *Journal of Marketing Research* 30, 1993, 305-314.

105. Thayer, M.A. "Contingent Valuation Techniques for Assessing Environmental Impacts: Further Evidence." *Journal of Environmental Economics and Management* 8, 1981, 27-44.

106. Thomas, J.F. and G.J. Syme. "Estimating Residential Price Elasticity of Demand for Water: A Contingent Valuation Approach." *Water Resources Research* 24, 1988, 1847-1857.

107. Vaughan, W.J. and C.S. Russell. *Freshwater Recreational Fishing: The National Benefits of Water Pollution Control.* Baltimore, MD: Resources for the Future/Johns Hopkins University Press, 1982.

108. Vaughan, W.J. and C.S. Russell. "Valuing a Fishing Day: An Application of a Systematic Varying Parameter Model." *Land Economics* 58, 1982, 450-463.

109. Varian, H.R. "Nonparametric Tests of Consumer Behavior." *Review of Economic Studies* 50, 1983, 99-110.

110. Walsh, R.G., D.M. Johnson and J.R. McKean. "Benefits Transfer of Outdoor Recreation Demand Studies: 1968-1988." *Water Resources Research* 28, 1992, 707-713.

111. Walsh, R.G., L.D. Sanders and J.B. Loomis. *Wild and Scenic River Economics: Recreation Use and Preservation Values*. Englewood, CO: American Wilderness Alliance, 1985.

112. Walsh, R.G., L.D. Sanders and J.R. McKean. "The Consumptive Value of Travel Time on Recreation Trips." *Journal of Travel Research* 29, 1990, 17-24.

113. Walsh, R.G., F.A. Ward and J.P. Olienyk. "Recreational Demand for Trees in National Forests." *Journal of Environmental Management* 28, 1989, 255-268.

114. Wegge, T.C., W.M. Hanemann and I.E. Strand. *An Economic Assessment of Marine Recreational Fishing in Southern California*, Report prepared for the National Marine Fisheries Service by Jones and Stokes Associates, Inc., Sacramento, CA, 1985.

115. Wegman, E.J. "Nonparametric Probability Density Estimation." *Technometrics* 14, 1972, 533-546.

116. White, H. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrics* 48, 1980, 817-838.

117. White, T.J. "A Comparison of Methods for Valuing Environmental Goods." Dissertation, *DAI* 50, 1989, 2584. State University of New York at Binghamton.

118. Willis, K.G. and G.D. Garrod. "Valuing Open Access Recreation on Inland Waterways." Countryside Change Working Paper No. 12. University of Newcastle, Newcastle, 1990.

119. Young, J.S. *et al. Net Economic Value of Upland Game Hunting in Idaho*, Resource Bulletin RM-15. Fort Collins, CO: U.S. Forest Service. Rocky Mountain Forest and Range Experiment Station, 1987.

FIGURE 1
COMPLETE DATASET: NONPARAMETRIC DENSITY ESTIMATE

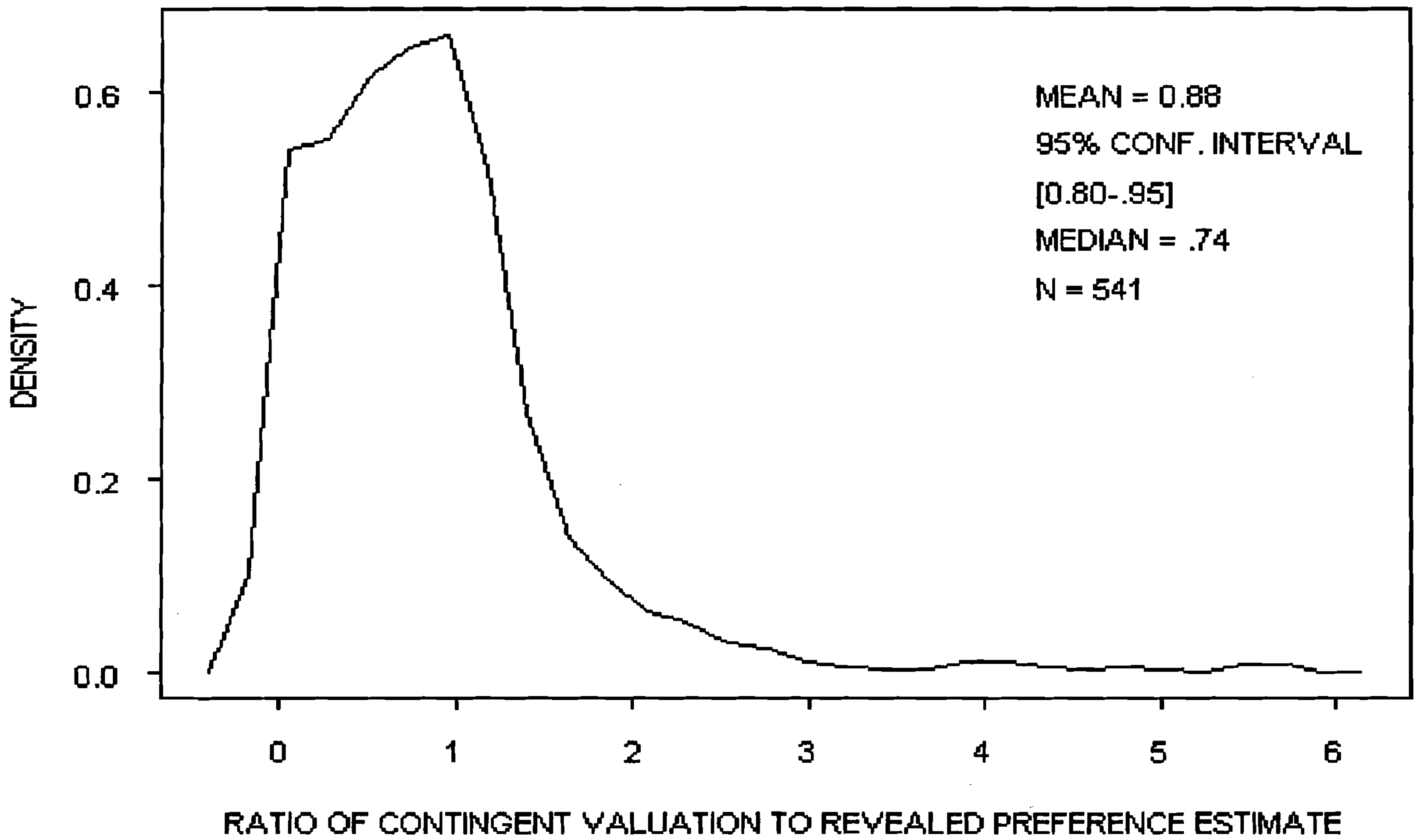


FIGURE 2
TRIMMED DATASET: NONPARAMETRIC DENSITY ESTIMATE

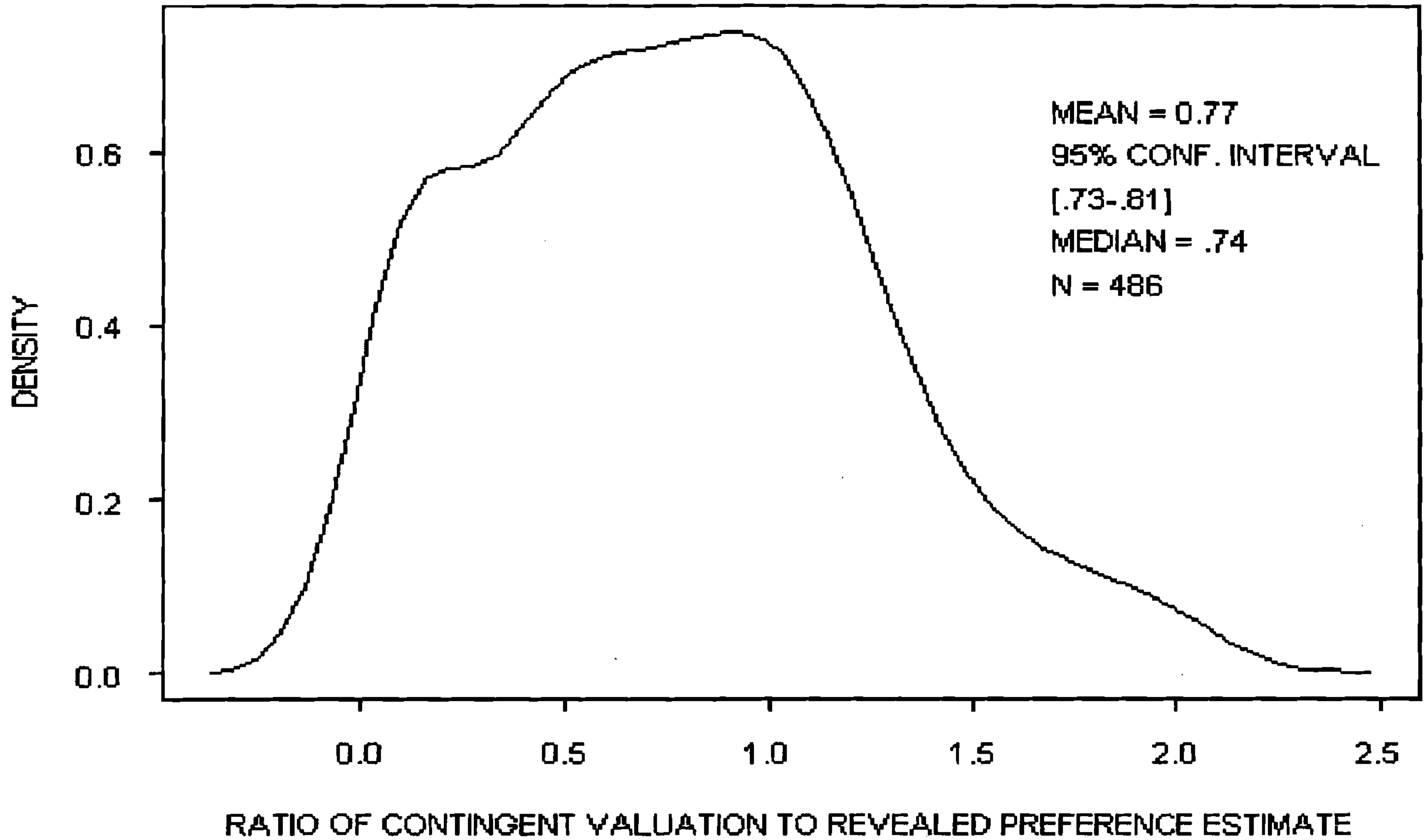
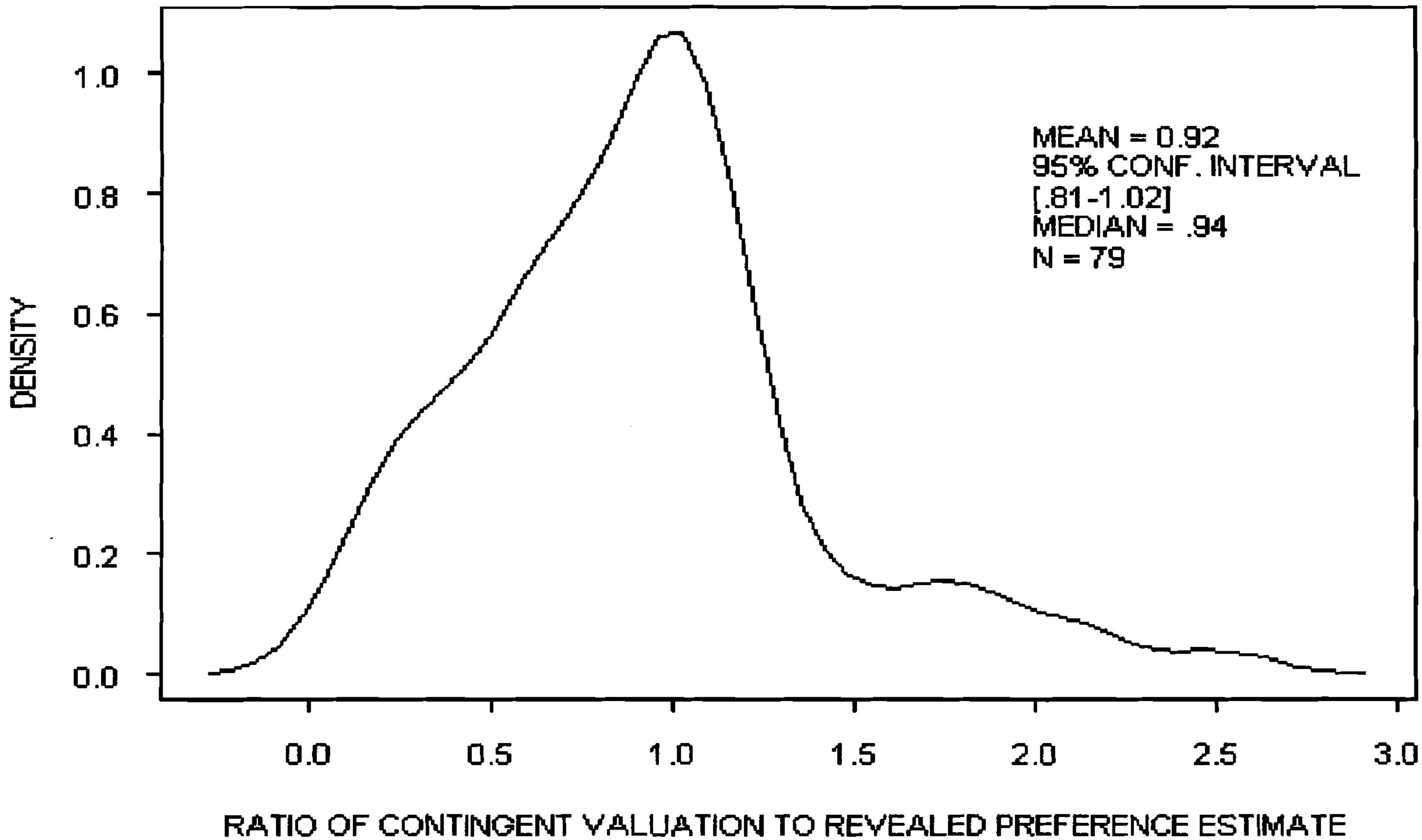


FIGURE 3

WEIGHTED DATASET: NONPARAMETRIC DENSITY ESTIMATE



SOME EVIDENCE CONCERNING THE VALIDITY OF CONTINGENT VALUATION:

PRELIMINARY RESULTS OF AN EXPERIMENT

by

Patricia A. Champ, Richard C. Bishop, Thomas C. Brown, and Daniel W. McCollum*

* Champ is a Graduate Student and Bishop is a Professor in the Department of Agricultural Economics at the University of Wisconsin-Madison. Brown and McCollum are economists at Rocky Mountain Forest and Range Experiment Station, U.S. Forest Service. The research was supported under a cooperative agreement between the College of Agricultural and Life Science, University of Wisconsin-Madison and the Rocky Mountain Forest and Range Experiment Station, Forest Service, USDA. Michael Hanemann and Richard Carson offered several helpful comments on an earlier draft.

SOME EVIDENCE CONCERNING THE VALIDITY OF CONTINGENT VALUATION: PRELIMINARY RESULTS OF AN EXPERIMENT

INTRODUCTION

The validity of the contingent valuation (CV) method to measure nonuse values is currently being debated. This research was undertaken in an effort to provide some evidence about whether CV can accurately measure the values citizens hold for an environmental amenity that has significant nonuse values. To explore this issue, we first define what is meant by "validity." Then we proceed to describe our research procedures. Finally, we assess whether there is evidence of CV providing a valid measure of the total value of a good with an large associated nonuse component.

VALIDITY

The American Psychological Association defines three kinds of validity - content, construct and criterion. A measure has content validity if it adequately covers the substance the measure is intended to cover. Content validity cannot be objectively assessed, rather it "... is a procedure which results in theoretical validity" (Bohrnstedt 1983, 97). Mitchell and Carson (1989) discuss two forms of construct validity -- convergent and theoretical. Tests of the convergent validity consider the relationship between the CV measure of a good's value and alternative measures of a good's value. Theoretical validity is assessed by considering the relationship between the CV measure and independent variables which are thought to be theoretically related to contingent values in predictable ways.

Criterion validity is defined as "the relation of the test to criteria outside the test itself" (Sundberg 1977, 44). To assess criterion validity, Mitchell and Carson (1989) say it is "necessary to have in hand a criterion which is unequivocally closer [than the contingent value] to the theoretical measure whose validity is being assessed" (Mitchell and Carson 1989, 192). In this study we developed and administered a survey instrument which we believe has content validity. The survey data were then used to assess both criterion validity and theoretical validity.[1] Specifically, a simulated market in which individuals were given the opportunity to actually pay for the good was implemented. This simulated market serves as the criterion against which the validity of the contingent values are assessed. Theoretical validity is assessed by considering the relationship between the contingent values and other measures which were elicited with the survey instrument. The following analysis is preliminary.

THE GOOD

The amenity valued in this experiment was carefully chosen to satisfy three conditions. First, there had to be a direct relationship between payment and provision. This relationship would help avoid individuals viewing the payment as a "donation" which may be motivated by reasons other than true willingness to pay (WTP). Second, the total value of the good needed to have a large nonuse value component. Previous validity experiments have dealt primarily with use values.[2] The National Oceanic and Atmospheric Administration report stressed the need for further research on the validity of CV for measuring nonuse values (Federal Register vol. 58, no. 10). Finally, the good had to be capable of actually being financed by the limited number of participants in this research project. After investigating several environmental projects, a road removal program in the North Rim of the Grand Canyon was selected.

The North Rim is open for visitation for only a few months each year. Although it is only ten miles from the South Rim as the crow flies, traveling between rims requires driving 215 miles by road. The North Rim is quite remote not only from the South Rim, but also from population centers in northern Arizona and southern Utah. As a result, very few visitors to Grand Canyon National Park visit the North Rim.

---

[1] We did not have the necessary data to assess the convergent form of construct validity.

[2] See Bohm 1972; Bishop and Heberlein 1980; Dickie, Fisher and Gerking 1987; Coursey, Hovis, and Schulze 1987; Kealy, Montgomery, and Dovidio 1990 for studies which assess the criterion validity of CV to measure use values.

Currently there are about 40 miles of compacted dirt roads in the North Rim that are no longer needed. The Grand Canyon National Park Service would like to remove these roads and designate the area as "wilderness." Volunteers are available who can remove the roads but no federal funding is allocated to pay for food and supplies for the volunteers. Food and supplies cost about $640 to remove one mile; each $1 removes eight feet of road.

STUDY PROCEDURES

In this study, the data were collected via mail questionnaires sent to a random sample of Wisconsin residents.[3] The mail procedures were fairly standard. First, an advance letter was sent to let people know that they would receive a questionnaire within a week. Then the survey materials, which included a cover letter, a question and answer sheet, and the questionnaires, were mailed. Each individual received two questionnaires: Part 1 and Part 2. Part 1 described the road removal program and contained a dichotomous choice format WTP question. Part 2 contained follow-up questions about why individuals responded as they did, national park experience, attitude questions, and demographic questions. A postcard and two follow-up mailings with questionnaires were sent to individuals who did not respond. There were two treatment groups: hypothetical market and simulated market. The two treatments received parallel survey materials except the cover letter to the simulated market group mentioned they would have a chance to actually pay for the resource and the Part 1 booklet for the simulated market group had the following paragraph after the WTP question: "If you decided to pay for the road removal program, please write a check to the 'Grand Canyon National Park Service' for the amount you said you are willing to pay and send the check when you return the two questionnaires (Part 1 and Part 2). We will forward all money to the Park Service at Grand Canyon National Park for the

[3] We did extensive preliminary design work including three focus groups, three pretests, and a pilot study to develop a survey instrument with content validity. We attempted to create a simulated market using a referendum. Unfortunately, we could not get the referendum to work. Based on the results of the pretests, we decided to use an individual dichotomous choice format for the pilot and final studies.

road removal program and they will send you a thank you note to acknowledge they received your payment." The Part 2 questionnaire was identical for both treatments.

RESPONSE RATES

The response rates were 51% for the hypothetical payment group and 44% for the actual payment group (see table 1). One reason the response rates were low may be that the topic is not highly salient to many Wisconsin residents. This is a very small project in a National Park that is far from Wisconsin. The difference in response rates between the hypothetical and simulated market treatments is statistically significant. However, the simulated and hypothetical market respondents were not significantly different on any of the demographic measures, experience with national parks or measures of environmentalism. Therefore, we are relatively confident in assuming that the two groups are representative of the same underlying population and any differences between the WTP of the two groups can be attributed to factors other than sampling or response rate effects.

IMPORTANT CAVEATS

Before we turn to the data, three caveats are in order. First, our experimental design is probably not fully incentive compatible. We set out originally to simulate a referendum, but failed to come up with a workable design. Hence, while our amenity almost certainly has public goods characteristics, we had to use an individual, voluntary payment mechanism, where participants only knew how much road would be removed for the dichotomous payment amount they themselves were asked to consider. However, in our defense, we would point out that it is not as obvious as it might appear at first glance what effects incentive compatibility ought to have. Our goal was to test how well CV does in predicting what people would actually pay. Since the CV scenario and the simulated market were set up in parallel fashion, to argue that our test of criterion validity was flawed, one would have to argue that the same flawed incentives had different effects in the CV exercise than they had in the simulated market. It is possible that this was the case, but no evidence exists one way or the other.

One might argue to the contrary that if CV works, participants in the CV exercise should have been able to predict how they would behaved under the posited incentives.

Our second caveat concerns the nature of the criterion we use to assess the validity of the contingent values. The simulated market required that participants who wanted to accept the opportunity to remove the specified number of feet of road at the specified price had to back up their acceptance by writing a check right away. Those who felt that they would like to accept, had not had much of an opportunity to anticipate this expenditure. Even if they failed to respond and as a result received the second and/or third follow-up surveys, nothing in earlier survey materials informed them that they would have such later opportunities. Only those who had the means at the time they filled out the survey could answer "yes" to the offer. By contrast, participants in the CV exercise were simply asked whether they would be willing to pay the specified amount. With the benefit of hindsight, we should have asked them whether they would be willing to write a check immediately. We did not. Thus, though we believe that the simulated market responses should still be interpreted as a criterion for purposes of validity testing, resulting values are perhaps best interpreted as a lower bound on true values.

Finally, while we did everything we could conceive of to try to legitimate the simulated market in the eyes potential participants, the link between their state university and Grand Canyon National Park may have remained unclear for some respondents. The effect of this problem, if it existed, would be to further encourage thinking of the simulated market values as lower bounds on true values.

RESULTS

As mentioned at the beginning of this paper, the purpose of this study was to assess two aspects of validity: theoretical and criterion. Assessments of theoretical validity for CV often involve estimating bid equations.[4] Criterion validity considers the relationship between hypothetical WTP and

[4] See Carson et al. 1992, Desvousges et al. 1992, and Boyce et al. 1992 for studies which assess the theoretical validity of the contingent valuation method.

actual WTP. The preliminary results of this study indicate hypothetical WTP may overstate actual WTP. However, hypothetical WTP appears to have theoretical validity.

Assessing Theoretical Validity

The preliminary analysis in this section looks at the relationships between the response to the WTP question and other variables. The results are based on contingency table analysis (Agresti, 1990). Logistic regression models were also estimated for each category of variables but the results are not reported unless they provide interesting information beyond that in the contingency table analysis. Each statistically tested proposition is stated as a hypothesis.

Hypothesis 1: There is an inverse relationship between the offer amount and responding positively to the WTP question.

Economic theory suggests an inverse relationship between the offer amount and the probability of saying "yes" to the WTP question.[5] In other words, as the offer amount increases fewer individuals should be willing to pay. Table 2 shows the percentage of respondents saying "yes" at each offer amount generally declines as the offer amount increases. The exception to this trend is the $8 and $12 interval for the hypothetical market treatment, but the difference between the percentage of respondents saying "yes" at $8 and at $12 is not significant. Likewise, for the simulated market respondents, the difference between 15% saying "yes" at $5 and 25% at $8 is not significant. Hence Hypothesis 1 is not rejected. The relationship between offer amount and the response to the WTP question is further analyzed by looking at the sign and significance of the coefficient on the offer variable in the simple logit model (table 3). The logit model is specified as:

[5] For the simulated market participants, saying "yes" means that they sent a check. There were six simulated market participants who said "yes" and did not send a check. Individuals who did not send a check were classified as saying "no" to the willingness to pay question. There were also individuals who sent a check for more than the amount they were asked to pay. These individuals were recorded as saying "yes" to the original offer amount (not the higher amount they paid).

$$P(yes) = \frac{1}{[1 + \exp^{\theta}]}$$

where offer is the amount the individual is asked to pay, P(yes) is the probability of a "yes" response to the offer, and $\Theta = -(\alpha + \beta^{*}\text{offer})$. The sign on the offer variable is negative as expected and the estimated coefficient is significantly different from zero.

Hypothesis 2: Individuals who are interested in the environment and contribute to environmental organizations are more likely to agree to pay for the road removal program.

One would expect variables which measure environmentalism to be positively related to WTP. Individuals who are more interested in the environment should be more willing to pay to preserve wilderness. Contingency table analysis confirms the expected relationship between WTP and interest in the environment for both the simulated and hypothetical market groups (see table 4). For the hypothetical market group, having contributed to an environmental organization makes a difference with respect to the response to the WTP question. Again the relationship is as expected, individuals who contributed in the past were more likely to say "yes" to the WTP question. The relationship does not hold for the simulated market respondents. This result may suggest individuals responding to the hypothetical payment situation draw on past experience to answer the WTP question whereas individuals who are actually paying may focus more on immediate factors such as the offer amount or their current finances.

Hypothesis 3: Individuals who have visited the Grand Canyon or expect to visit the Grand Canyon in the future are more likely to pay for the road programs.

The variables which measure national park experience that have a significant relationship to the response to the WTP question include having ever visited a national park (individuals who visited a national park were more likely to say "yes") and expectations about visiting Grand Canyon National

Park in the future (individuals who think they will visit in the future are more likely to say "yes"). This expectation is supported for both the hypothetical and simulated market respondents (see table 5). We cannot reject Hypothesis 3.

Hypothesis 4: Pro-environment attitudes are positively related to WTP.

Respondents were asked whether they disagree strongly, disagree, neither disagree nor agree, agree, or agree strongly with eleven attitude statements. These statements were designed to measure aspects of the underlying preferences for wilderness areas and environmental amenities. The responses to these attitude variables were distributed similarly for the hypothetical and simulated market groups. Likewise, for both groups, the attitude variables were related significantly to how the individual responded to the WTP question (see table 6). As expected, individuals with pro-environment attitudes were more likely to pay for the road removal program. For example, 67 percent of the respondents who said "yes" to hypothetical WTP question said that they agreed strongly with the statement: "I would like for wilderness areas to be preserved even if I never get to visit them." By comparison, only 32 percent of the respondents who said "no" to the hypothetical WTP question said they agreed strongly with this statement. The significant relationship of the attitude variables to both hypothetical and actual WTP suggests that CV does reflect underlying preferences.

Hypothesis 5: An individual's socioeconomic background affects WTP.

The demographic variables are related to the response to the WTP question as economic theory would predict. The individual's area of residence, age, education, and income all have significant relationships with the response to the WTP question for both hypothetical and simulated market respondents. Specifically, individuals from urban areas were more likely to say "yes" to the WTP question than individuals from rural areas. Likewise, individuals who said "yes" to the WTP question were more educated and had higher incomes than individuals who said "no." Gender and having children did not seem to influence the response to the WTP question (see table 7).

In summary, the relationships between the construct of interest (response to WTP question) and variables which economic theory would predict to be related to the construct suggest there is evidence of theoretical validity. This relationship can be further explored by estimating a bid equation with relevant variables (table 8). The estimated bid equation includes eight independent variables: offer, one statement variable about whether the individual would rather see the money go to a better project, five attitude variables, and income. Since some of these demographic variables are interrelated, only income is included in the estimated bid equation. The effect, as reflected in the magnitude of the estimated coefficient, of the offer amount on response is greater for individuals who actually pay relative to those who hypothetically pay. Individuals whose payment is hypothetical seem to draw more on their attitudes toward wilderness when responding to the WTP question. The significant relationship between willingness to pay and the offer, attitudes, and income are positive evidence of the theoretical validity.

Assessing Criterion Validity

Relationships are again analyzed using simple contingency table analysis. Given that the two groups were randomly selected from the same population, one would expect the two groups to respond in a similar manner to objective or factual questions. The nature of the WTP exercise (actual or hypothetical payment) may have an effect on how the individual perceives and responds to the WTP question.

Hypothesis 6: Hypothetical WTP is the same as actual WTP.

To test this hypothesis, we look at the response to the WTP question at the various offer amounts (table 2). At five of the six offer amounts, a significantly greater percentage of respondents in the hypothetical group (than the simulated group) said "yes" they would pay the amount asked about. At $8, the difference between the percentage of hypothetical respondents and the simulated market respondents who said "yes" is not significant. Another way to look at the relationship between the response to the offer amount and the treatment group is to consider the WTP function (table 3). If a separate logit model is estimated for each treatment group, the hypothesis that the coefficients are equal can be tested. The results of estimating simple logit models are shown in table 3. The null hypothesis that the hypothetical and simulated market distributions are the same can be tested (H_o: $\alpha_{hyp} = \alpha_{sim}$, $\beta_{hyp} = \beta_{sim}$) using the likelihood ratio test.

$$LR = -2*[(LL_{hyp} + LL_{sim}) - LL_{full}] \sim \chi^2\ (r)$$

where LL_{hyp} and LL_{sim} are the log likelihood values associated respectively with the hypothetical and the simulated treatment models, LL_{full} is the log likelihood value associated with the model using both the hypothetical and simulated market data sets, and r is the number of restrictions imposed. The likelihood ratio test suggests that the hypothetical and simulated market models are distributed differently at the five percent level (LR = 69.61 > $\chi_{2,5}$ = 5.99). Therefore it appears as though individuals respond differently to the offer amount depending on whether payment is real or hypothetical; more individuals say "yes" when the payment is hypothetical. The estimated mean WTP measures are shown in table 9. The difference between the mean WTP for the two treatments is significant. The estimated distribution for the hypothetical market data is very crude since there are no observations in the upper tail (at the top offer amount of $50, 34% of the hypothetical market respondents said yes). Given the poor fit of the hypothetical market WTP distribution, the mean estimate is suspect as is the magnitude of the difference between the two means. Note in particular the huge 95 percent confidence interval around the estimated mean from the CV data.

Many explanations for the difference between CV responses and actual payment have been put forth. Strategic behavior, warm glow effects, failure to consider income constraints, and difficulties in dealing with hypothetical questions are only a few examples. There seems to be a tendency to assume that all respondents to the CV question make the same "mistake" whatever the mistake may be. Perhaps some individuals answer correctly and other individuals make mistakes in the sense that they answer "yes" to dichotomous choice questions when they would not in fact pay or "no" when they would. We would propose the following procedures for attempting to identify which respondents made "mistakes."

Particularly given the caveats stated earlier, the term "mistake" is to some extent a convenience. A respondent has made a mistake as we shall use the term when she or he answers "yes" to a CV question while we predict based on statistical analysis that he or she would have said "no" in the simulated market or vice versa.

WHO MADE MISTAKES?

Procedures

Our proposed procedures for defining who in the hypothetical payment group made a "mistake" are as follows. First, using the simulated market data, we estimated a WTP function with responses to the WTP question as the dependent variable and offer, one statement variable about whether the individual would rather see the money go to a better project, five attitude variables, and income as the independent variables. For purposes of analysis, this model is considered the "true" WTP model since individuals who said "yes" to the WTP question had actually sent a check. See table 8 for a description of the variables and the estimated model. Note in particular that the model fitted to the simulated market data predicts correctly whether a simulated market participant will respond "yes" or "no" 86 percent of the time. Next, the coefficient estimates from the "true" model were used to predict the probability of a hypothetical respondent actually paying. Individuals with a probability greater than .5 were coded as saying "yes" to actual WTP and all others were coded as saying "no" to an actual WTP question.[6] Finally, the predicted response to an actual WTP question is compared to the actual response to the hypothetical WTP question for each individual in the hypothetical market group. Responses of individuals who said "yes" to the hypothetical WTP question and were predicted as saying "no" to the actual payment question were classified as "mistakes." Likewise responses of individuals who said "no"

---

[6] Of course, criteria other than .5 could have been used. The actual responses to the CV questions over all offer amounts were 44 percent Yes and 56 percent No. For the simulated market the comparable figures are 17 percent and 83 percent respectively. When the model was fitted to the simulated data and then applied to the CV respondents using the .5 criterion, it predicted 18 percent Yes and 82 percent No. This appears to support using a predicted probability of saying Yes at .5 as the criterion for setting the predicted responses at Yes.

to the hypothetical WTP question and were predicted to say "yes" to the actual payment question were also classified "mistakes." Responses of individuals with the same predicted actual payment and hypothetical WTP were classified as "non-mistakes."[7]

Results

As shown in table 10, 35 percent of the respondents to the hypothetical WTP question made a mistake.[8] Only six individuals (about two percent) made a mistake in the sense of saying "no" to the hypothetical question when it is predicted they would say "yes" if payment were real. The rest of the respondents who made a mistake said "yes" to the hypothetical WTP when it is predicted they would not actually pay.

The next interesting question is whether the individuals who made mistakes differ from those who did not. Since 95 percent of the people who made a mistake said "yes" to the hypothetical WTP question, the following analysis compares the mistakes to the non-mistakes for those who said "yes" to the CV question.

The offer amount seems to have affected whether an individual makes a mistake (table 11). Respondents made more mistakes as the offer amount increased.

The mistakes and non-mistakes seem to have said "yes" for the some of the same reasons (table 12). However, non-mistakes were more likely to say that they definitely agree with the statement "I felt that the road removal project would be worth the amount I decided to pay." Likewise non-mistakes were more likely to say they definitely agree with the statement "Restoring a wilderness area is very important to me." Seventy-eight percent of the non-mistakes said that they definitely agree with this statement compared to only 58 percent of the mistakes.

---

[7] Predicted responses could only be estimated for the cases that did not have missing data on any of the variables used in the model. The original hypothetical payment group had 393 cases but only 316 of the cases could be used for this analysis.

[8] It should be recalled that with the simulated market data, 85 percent of the observed responses are equal to the predicted responses. So, while the model does predict rather well, it is not "perfect."

Past experience with a National Park seemed to affect whether an individual makes a mistake when responding to the CV question. Ninety percent of the non-mistakes compared to 71 percent of the mistakes had visited a National Park in the past (table 13). Likewise, the non-mistakes are more likely than mistakes to say they will visit the North Rim in the future. Twenty-four percent of the non-mistakes compared to nine percent of the mistakes said it is very likely they will visit the North Rim in the future.

The mistakes and non-mistakes have different attitudes toward National Parks and the environment in general. The pattern seems to be that the mistakes are more ambivalent about the attitude statements and the those who did not make mistakes tend to be distributed at the extremes of the attitude scales; they either disagree or agree strongly with the attitude statements. Table 14 shows the distribution of responses to the attitude variables for both the mistakes and the non-mistakes.

OTHER ISSUES

On a broader level, some argue that contingent values do not measure underlying preferences (Diamond and Hausman, 1992). The results of this study suggest that responses to the hypothetical WTP question are consistent with preferences as measured by individuals' attitudes toward wilderness. It has also been argued that hypothetical WTP reflects "the willingness to pay for moral satisfaction of contributing to public goods, not the economic value of these goods" (Kahneman and Knetsch 1992, 57). However, in this study 68 percent of the respondents who said "yes" to the hypothetical WTP question said it was "definitely true" and 32 percent said it was "somewhat true" that the road removal program was worth the amount they said they would pay. This result suggests that most of the individuals answering the hypothetical WTP question were considering the value of the program to them. There was some evidence of individuals saying "yes" for more general reasons than the value of the program but that was not the case for the majority of the respondents. A final criticism of using CV to measure nonuse values gets at the foundation of the concept. Some argue that individuals do not receive positive benefits from an environmental improvement in an area the individual is not familiar with and never plans to use. Yet in this study, 96 percent of the people who sent a check had never been to the North Rim and 45 percent said it was "very unlikely" that they will ever visit the area. Of course, users of the resource may also have a large associated nonuse value.

CONCLUSIONS

Does CV provide a "valid" measure of true WTP? Unfortunately, based on the study reported here we cannot offer a clear "yes" or "no." The responses to the CV question are related to other explanatory variables in significant ways as economic theory would predict. Specifically, individuals with positive attitudes toward wilderness areas and higher incomes were more likely to say yes. On the other hand, the higher the offer amount the more likely individuals were to say "no" to the WTP question. Based on these observations, it appears as though the CV method has theoretical validity. This research also allowed for a test of criterion validity. Simulated market values are assumed to be more closely related (than hypothetical values) to the underlying true WTP and serve as the criterion. Statistical tests of criterion validity suggest that the contingent values are significantly higher than simulated market values. Unfortunately, the data from this study do not provide much insight into the magnitude of the difference between hypothetical and actual WTP. Instead, a "fat" right hand tail leaves a CV estimate that is to unreliable to serve this function. Furthermore, questions can be raised about the adequacy of our criterion. The lack of incentive compatibility, the requirement that responding positively in the simulated market required unforeseen commitments, and other concerns may mean that our simulated market underestimated true values. The analysis of the "mistakes" versus the "non-mistakes" in the hypothetical payment group suggests that the majority (65%) of the respondents to our CV question did reveal their true preferences. If an experiment could be designed to avoid some of the pitfalls we encountered, the gap we found between CV and simulated market values might be narrowed even further or even disappear entirely. Given the potential importance of nonuse values and the controversy that CV applications have stimulated, the need for more research is obvious.

Table 1: Response Rate by Treatment Group

	Hypothetical Payment Group	Actual Payment Group
Initial Sample	850	850
Undeliverable	85	82
Completed surveys	392	340
Response rate	51%	44%

Note: Response Rate = Complete Surveys/(Initial Sample - Undeliverable).

Table 2: Percentage of Hypothetical and Simulated Market Respondents Who Said "Yes" to the WTP Question

Offer Amount	Hypothetical market respondents who said "yes"	Simulated market respondents who said "yes"
$1*	53%	24%
$5*	51%	15%
$8	39%	25%
$12*	48%	17%
$15*	39%	13%
$50*	34%	4%

* Using the test of equality of two population proportions the difference between the percentage of hypothetical and simulated market respondents who said "yes" to the WTP question is significant at the 5% level for this offer amount.

Table 3: Estimated Logit Equation

	Hypothetical Market Group	Simulated Market Group	Combined Data Set
Constant	-0.039	-1.135**	-0.536**
(Std. Error)	(0.139)	(0.198)	(0.107)
Offer	-0.014**	-0.040**	-0.018**
(Std. Error)	(0.007)	(0.015)	(0.006)
Log Likelihood	-267.32	-149.93	-451.06
Number of observations	393	339	732

** Significant at 5% level.

Table 4: Environmentalism of Respondents

Variable	Hypothetical Market Respondents		Simulated Market Respondents	
	"Yes" to WTP Question	"No" to WTP Question	"Yes" to WTP Question	"No" to WTP Question
Interest in the environment:[1, 2]				
Not at all interested in the environment	0%	4%	0%	3%
A little interested in the environment	6%	12%	2%	12%
Interested in the environment	36%	49%	33%	49%
Strongly interested in the environment	43%	28%	44%	26%
Extremely interested in the environment	15%	7%	21%	10%
Contributed money or volunteered for an environmental organization:[1]				
Yes	44%	18%	30%	23%
No	56%	82%	70%	77%

[1] Distributions of responses for individuals who said "yes" and individuals who said "no" to the WTP question are significantly different at 5% level for hypothetical market respondents.

[2] Distributions of responses for individuals who said "yes" and individuals who said "no" to the WTP question are significantly different at 5% level for simulated market respondents.

Table 5: Grand Canyon Experience

Variable	Hypothetical Market Respondents		Simulated Market Respondents	
	"Yes" to WTP Question	"No" to WTP Question	"Yes" to WTP Question	"No" to WTP Question
Ever visit a National Park in the United States:[1,2]				
No	22%	35%	16%	31%
Yes	78%	65%	84%	69%
Ever visit Grand Canyon National Park (for individuals who visited a National Park):				
No	62%	56%	55%	59%
Yes	38%	44%	45%	41%
Ever visit the South Rim of the Grand Canyon (for individuals who visited the Grand Canyon):				
No	6%	7%	14%	6%
Yes	94%	93%	86%	94%
Ever visit the North Rim of the Grand Canyon (for individuals who visited the Grand Canyon):				
No	78%	76%	96%	85%
Yes	22%	24%	4%	15%
Ever hike the Grand Canyon (for individuals who visited the Grand Canyon):				
No	68%	76%	77%	76%
Yes	32%	24%	23%	24%
How likely visit North Rim in the future:[1]				
Very unlikely	32%	49%	45%	56%
Somewhat unlikely	18%	19%	15%	17%
Somewhat likely	37%	22%	26%	20%
Very likely	13%	10%	13%	7%

Table 5: Grand Canyon Experience

Variable	Hypothetical Market Respondents		Simulated Market Respondents	
	"Yes" to WTP Question	"No" to WTP Question	"Yes" to WTP Question	"No" to WTP Question
How likely visit Grand Canyon National Park in the future:[1, 2]				
Very unlikely	18%	35%	20%	35%
Somewhat unlikely	20%	17%	9%	17%
Somewhat likely	34%	29%	32%	30%
Very likely	28%	19%	39%	18%

1 Distributions of responses for individuals who said "yes" and individuals who said "no" to the WTP question are significantly different at 5% level for hypothetical market respondents.

2 Distributions of responses for individuals who said "yes" and individuals who said "no" to the WTP question are significantly different at 5% level for simulated market respondents.

Table 6: Distribution of Responses to Attitude Statements

Statement	Group and response to WTP question		Disagree Strongly	Disagree	Neither Disagree nor Agree	Agree	Agree Strongly
All areas of National Parks should be easily accessible by roads	HYP:[1]	Yes	43%	39%	8%	6%	4%
		No	17%	38%	21%	18%	6%
	SIM:[1]	Yes	40%	42%	7%	5%	5%
		No	19%	41%	21%	12%	7%
It is important to me that future generations be able to enjoy wilderness areas	HYP:[1]	Yes	1%	1%	2%	25%	72%
		No	2%	1%	7%	52%	38%
	SIM:[1]	Yes	0%	0%	2%	28%	70%
		No	3%	2%	5%	51%	39%
I would like for wilderness areas to be preserved even if I never get to visit them	HYP:[1]	Yes	1%	0%	2%	31%	67%
		No	2%	3%	13%	49%	32%
	SIM:[1]	Yes	0%	0%	0%	32%	68%
		No	3%	1%	7%	48%	40%
Funding for protection of wilderness areas should come primarily from state and federal governments instead of private donations	HYP:	Yes	3%	18%	36%	31%	13%
		No	5%	16%	31%	32%	16%
	SIM:	Yes	2%	21%	40%	32%	5%
		No	7%	16%	32%	29%	17%
I think it is everyone's responsibility to help the environment any way we can	HYP:[1]	Yes	1%	2%	4%	47%	47%
		No	2%	5%	19%	54%	20%
	SIM:[1]	Yes	0%	0%	2%	40%	58%
		No	3%	4%	13%	53%	27%
I care about wilderness areas outside of Wisconsin	HYP:[1]	Yes	1%	0%	2%	51%	46%
		No	1%	1%	15%	60%	22%
	SIM:[1]	Yes	0%	0%	2%	51%	47%
		No	2%	2%	15%	58%	24%

Table 6: **Distribution of Responses to Attitude Statements**

Statement	Group and response to WTP question		Disagree Strongly	Disagree	Neither Disagree nor Agree	Agree	Agree Strongly
National parks should be managed to preserve native species	HYP:[1]	Yes	0%	2%	6%	44%	49%
		No	0%	3%	14%	58%	25%
	SIM:[1]	Yes	0%	2%	2%	44%	53%
		No	0%	2%	13%	55%	30%
National parks should be managed to preserve wilderness areas	HYP:[1]	Yes	0%	0%	5%	49%	46%
		No	0%	5%	17%	58%	19%
	SIM:[1]	Yes	0%	0%	2%	39%	60%
		No	1%	2%	12%	57%	28%
I think too much federal and state money is spent on wilderness preservation programs	HYP:[1]	Yes	22%	47%	29%	1%	1%
		No	10%	32%	41%	14%	4%
	SIM:[1]	Yes	37%	30%	33%	0%	0%
		No	14%	28%	44%	10%	4%
As much wilderness as possible should be preserved no matter what the cost	HYP:[1]	Yes	8%	27%	32%	24%	9%
		No	23%	36%	25%	12%	4%
	SIM:[1]	Yes	4%	25%	23%	35%	14%
		No	13%	36%	27%	16%	7%
My responses to this study will be important in future decisions about the environment	HYP:[1]	Yes	1%	10%	44%	36%	9%
		No	6%	13%	50%	27%	4%
	SIM:[1]	Yes	0%	6%	43%	38%	13%
		No	6%	13%	53%	25%	4%

Note: SIM refers to the simulated market group and HYP refers to the hypothetical market group.

1 Distributions of responses for "yes" and "no" respondents to the WTP question are significantly different at 5% level.

Table 7: Demographic Background of Simulated and Hypothetical Market Respondents

Variable	Hypothetical Market Respondents		Simulated Market Respondents	
	"Yes" to WTP Question	"No" to WTP Question	"Yes" to WTP Question	"No" to WTP Question
Area of residence:[1]				
Urban	44%	34%	60%	32%
Suburban	21%	15%	12%	20%
Rural	34%	51%	28%	49%
Have children:				
No	23%	22%	32%	22%
Yes	77%	78%	68%	78%
Age:[1]				
Less than 30	8%	6%	14%	7%
30 to 39	28%	21%	16%	21%
40 to 49	26%	19%	21%	25%
50 to 59	15%	22%	18%	15%
60 to 69	15%	14%	16%	18%
More than 69	8%	19%	14%	14%
Gender:				
Male	70%	77%	75%	75%
Female	30%	23%	25%	25%
Education:[1]				
Eighth grade or less	2%	7%	0%	6%
Some high school	2%	5%	4%	6%
High school graduate	16%	23%	19%	25%
Some college or technical school	24%	26%	23%	24%
Technical or trade school graduate	14%	9%	5%	12%
College graduate	19%	14%	23%	14%
Some graduate work	9%	9%	7%	6%
Advanced degree	13%	8%	19%	8%

Table 7: Demographic Background of Simulated and Hypothetical Market Respondents

Variable	Hypothetical Market Respondents		Simulated Market Respondents	
	"Yes" to WTP Question	"No" to WTP Question	"Yes" to WTP Question	"No" to WTP Question
Income:[1]				
Less than $10,000	4%	10%	8%	12%
$10,000 to $19,999	10%	16%	9%	15%
$20,000 to $29,999	16%	20%	15%	21%
$30,000 to $39,999	14%	22%	15%	18%
$40,000 to $49,999	10%	11%	11%	10%
$50,000 to $59,999	16%	10%	13%	10%
$60,000 to $69,999	8%	5%	11%	6%
$70,000 to $79,999	6%	2%	2%	3%
$80,000 to $99,999	6%	2%	4%	4%
$100,000 or more	10%	2%	11%	1%

[1] Distributions of responses for "yes" and "no" to WTP question are significantly different at 5% level.

Table 8: Estimated Logit Model for Full Model

Description of Independent Variables	HYP Estimated Coefficients (Std. Error)	SIM Estimated Coefficients (Std. Error)	Estimated Coefficients for Combined Model (Std. Error)
Constant	-10.563** (1.564)	-12.892** (2.419)	-10.275** (1.14)
Offer	-.029** (.010)	-.050** (.020)	-.026** (.008)
I would rather see the money go to a better project[1]	1.283** (.183)	1.182** (.248)	1.084** (.128)
I would like for wilderness areas to be preserved even if I never get to visit them[2]	.487* (.256)	.4426 (.498)	.448** (.206)
I think it is everyone's responsibility to help the environment any way we can[2]	.289 (.220)	.845** (.423)	.288 (.178)
I care about wilderness areas outside of Wisconsin[2]	.344 (.312)	-.293 (.367)	.131 (.211)
As much wilderness as possible should be preserved no matter what the cost[2]	.273* (.154)	.141 (.215)	.175 (.113)
My responses to this study will be important in future decisions about the environment[2]	.182 (.209)	.726** (.268)	.461** (.148)
Income	.325** (.071)	.324** (.093)	.307** (.051)
Log Likelihood	-136.87	-77.74	-249.63
Percent of Responses Predicted Correctly	80.06%	85.89%	77.38%
Number of Observations	316	241	557

** Significant at 5% level.

[1] This statement was ranked on a four point scale: 1= definitely true, 2=somewhat true, 3=somewhat false, 4=definitely false.

[2] This statement was ranked on a five point scale: 1=disagree strongly, 2=disagree, 3=neither disagree nor agree, 4=agree, 5=agree strongly.

Table 9:	Estimated mean WTP for hypothetical and simulated market respondents	
	Mean[1]	95% Confidence Interval
Hypothetical Market	46.33	[26.66, 283.78]
Simulated Market	7.20	[4.65, 23.54]

[1] Mean was estimated using Hanemann formula (Hanemann 1989).

Table 10: Percentage of Mistakes and Non-mistakes (N=316) for Respondents in the Hypothetical Treatment Group

Non-mistakes	65%
Mistakes who said "no" to hypothetical WTP but predict "yes" to actual WTP	2%
Mistakes who said "yes" to hypothetical WTP but predict "no" to actual WTP	33%

Table 11: Distribution of Offer Amounts for Mistakes and Non-Mistakes Who Said "Yes" to the WTP Question

Offer Amount	Non-mistakes	Mistakes
$1	48%	52%
$5	48%	52%
$8	23%	77%
$12	24%	76%
$15	24%	76%
$50	12%	88%

Table 12: Distribution of Responses to Statements about "Yes" Response to WTP Question -- Mistakes versus Non-mistakes

Statement	Group	Definitely True	Somewhat True	Somewhat False	Definitely False
I felt that the road removal project would be worth the amount I decided to pay[1]	NMIS	90%	8%	0%	2%
	MIS	57%	43%	0%	0%
I decided to pay more to show my support for the environment in general rather than specifically for the road removal program	NMIS	46%	29%	15%	10%
	MIS	29%	31%	28%	12%
The total number of feet of road that would be removed was important in my decision to pay for the road removal program	NMIS	16%	43%	16%	25%
	MIS	11%	32%	30%	27%
Restoring a wilderness area is very important to me[1]	NMIS	78%	22%	0%	0%
	MIS	58%	41%	0%	0%

Note: NMIS refers to the non-mistake group and MIS refers to the mistake group.

1 Distributions of responses for NMIS and MIS are significantly different at 5% level.

Table 13: **Past and Expected Future Grand Canyon Experience**

Variable	Non-Mistakes	Mistakes
Ever visit a National Park in the United States:[1]		
No	14%	38%
Yes	86%	62%
Ever visit Grand Canyon National Park (for individuals who visited a National Park):		
No	71%	58%
Yes	29%	42%
Ever visit the South Rim of the Grand Canyon (for individuals who visited the Grand Canyon):		
No	8%	3%
Yes	93%	97%
Ever visit the North Rim of the Grand Canyon (for individuals who visited the Grand Canyon):		
No	77%	77%
Yes	23%	23%
Ever hike or backpack in Grand Canyon National Park (for individuals who visited the Grand Canyon):		
No	54%	71%
Yes	46%	29%
How likely visit North Rim in the future:[1]		
Very unlikely	9%	42%
Somewhat unlikely	13%	19%
Somewhat likely	54%	31%
Very likely	24%	9%
How likely visit Grand Canyon National Park in the future:[1]		
Very unlikely	8%	21%
Somewhat unlikely	14%	24%
Somewhat likely	39%	31%
Very likely	39%	24%

[1] Distributions of responses for NMIS and MIS are significantly different at the 5% level.

Table 14: Distribution of Responses to Attitude Statements -- Non-mistakes versus Mistakes

Statement	Group	Disagree Strongly	Disagree	Neither Disagree nor Agree	Agree	Agree Strongly
All areas of National Parks should be easily accessible by roads[1]	NMIS	53%	39%	2%	2%	4%
	MIS	39%	37%	11%	10%	3%
It is important to me that future generations be able to enjoy wilderness areas[2]	NMIS	2%	0%	0%	8%	90%
	MIS	0%	0%	4%	31%	65%
I would like for wilderness areas to be preserved even if I never get to visit them[2]	NMIS	2%	0%	0%	12%	86%
	MIS	0%	0%	2%	36%	62%
Funding for protection of wilderness areas should come primarily from state and federal governments instead of private donations	NMIS	4%	14%	41%	28%	14%
	MIS	3%	22%	32%	34%	9%
I think it is everyone's responsibility to help the environment any way we can[2]	NMIS	0%	0%	2%	26%	72%
	MIS	1%	3%	6%	54%	37%
I care about wilderness areas outside of Wisconsin[2]	NMIS	0%	0%	2%	35%	63%
	MIS	0%	0%	3%	58%	39%
National parks should be managed to preserve native species[2]	NMIS	0%	4%	4%	26%	67%
	MIS	0%	1%	8%	50%	42%
National parks should be managed to preserve wilderness areas[2]	NMIS	0%	0%	6%	33%	61%
	MIS	0%	0%	5%	54%	42%

Table 14: **Distribution of Responses to Attitude Statements -- Non-mistakes versus Mistakes**

Statement	Group	Disagree Strongly	Disagree	Neither Disagree nor Agree	Agree	Agree Strongly
I think too much federal and state money is spent on wilderness preservation programs[1]	NMIS	31%	47%	20%	0%	2%
	MIS	16%	48%	35%	1%	0%
As much wilderness as possible should be preserved no matter what the cost[2]	NMIS	6%	14%	31%	33%	16%
	MIS	9%	32%	33%	20%	7%
My responses to this study will be important in future decisions about the environment[2]	NMIS	0%	6%	33%	33%	28%
	MIS	1%	13%	50%	34%	1%

Note: NMIS refers to the non-mistake group and MIS refers to the mistake group.

1 Distributions of responses for NMIS and MIS are significantly different at the 10% level.

2 Distributions of responses for NMIS and MIS are significantly different at the 5% level.

BIBLIOGRAPHY

Agresti, A. 1990. *Categorical Data Analysis.* John Wiley and Sons, New York.

Bishop, R.C. and T.A. Heberlein, 1980. "Simulated Markets, Hypothetical Markets, and Travel Cost Analysis: Alternative Methods of Estimating Outdoor Recreation Demand", Staff Paper Series no. 187, Department of Agricultural Economics, University of Wisconsin.

Bohm, P., 1972. "Estimating Demand for Public Goods: An Experiment", *European Economic Review*, 3:111-130.

Bohrnstedt, George W. 1983. "Measurement." In *Handbook of Survey Research.* Eds. P.H. Rossi, J.D. Wright, and A.B Anderson. New York: Academic Press, Inc.

Boyce, R.R., T.C. Brown, G.D. McClelland, G.L. Peterson, and W.D. Schulze, 1989. "Experimental Evidence of Existence Value in Payment and Compensation Contexts", Paper presented at the joint meetings of the Western Committee on the Benefits and Costs of Natural Resource Planning (W133) and the Western Regional Science Association, San Diego, CA.

Carson, Richard T., Robert C. Mitchell, W. Michael Hanemann, Raymond J. Kopp, Stanley Presser, and Paul A. Ruud. 1992. "A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill." A Report Submitted to the Attorney General of the State of Alaska. Reprinted by Natural Resource Damage Assessment, Inc.

Coursey, D.L., J.J. Hovis, and W.D. Schulze, 1987. "The Disparity between Willingness to Accept and Willingness to Pay Measures of Value", *Quarterly Journal of Economics*, 102:679-90.

Desvousges, W.H., F.R. Johnson, R.W. Dunford, K.J. Boyle, S.P. Hudson, and K.N. Wilson, 1992. "Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability", paper presented at the Cambridge Economics, Inc. symposium titled "Contingent Valuation: A Critical Assessment", Washington, D.C., April 2 and 3.

Diamond, P.A. and J.A. Hausman, 1992. "On Contingent Valuation Measurement of Nonuse Values", paper presented at the Cambridge Economics, Inc. symposium titled "Contingent Valuation: A Critical Assessment", Washington, D.C., April 2 and 3.

Dickie, M., A. Fisher, and S. Gerking, 1987. "Market Transactions and Hypothetical Demand Data: A Comparative Study", *Journal of the American Statistical Association*, 82:69-75.

Hanemann, W.M., 1989. "Welfare Evaluation in Contingent Valuation Experiments with Discrete Responses: Reply", *American Journal of Agricultural Economics*, 71(4):1057-1061.

Kahneman, D. and J.L. Knetsch, 1992. "Valuing Public Goods: The Purchase of Moral Satisfaction." *Journal of Environmental Economics and Management*, 22(1):57-70.

Kealy, M.J., M. Montgomery, J.F. Dovidio, 1990. "Reliability and Predictive Validity of Contingent Values: Does the Nature of the Good Matter?", *Journal of Environmental Economics and Management*, 19:244-263.

Mitchell, R.C., and R.T. Carson, 1989b. *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Resources for the Future, Washington, D.C.

Seip, K. and J. Strand, 1992. "Willingness to Pay for Environmental Goods in Norway: A Contingent Valuation Study with Real Payment." *Environmental and Resource Economics* 2(1):91-106.

Sundberg, N.D., 1977. *Assessment of Persons*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

U.S. Department of Commerce. National Oceanic and Atmospheric Administration. 1993. *Natural Resource Damage Assessments under the Oil Pollution Act of 1990.* Federal Register 58(10):4601-4614.

STATISTICAL AND PSYCHOLOGICAL INFLUENCES ON CONTINGENT VALUATION WILLINGNESS TO PAY ESTIMATES

by

Jean C. Buzby, Richard C. Ready, and Dayuan Hu*

ABSTRACT

Willingness to pay for food safety improvements estimated using the payment card elicitation technique was smaller than that using dichotomous choice. A small part of this difference was due to bias introduced by the statistical techniques. Most of the difference was due to differences in respondent behavior.

* Authors are Research Specialist, Assistant Professor, and Ph.D. Candidate in the Department of Agricultural Economics at the University of Kentucky. The views expressed in this article are the authors' and do not necessarily represent the views of the University of Kentucky.

STATISTICAL AND PSYCHOLOGICAL INFLUENCES ON CONTINGENT VALUATION WILLINGNESS TO PAY ESTIMATES[1]

INTRODUCTION

The contingent valuation (CV) method has become increasingly used to measure values associated with a variety of nonmarket goods such as water quality, toxic waste dumps, recreation, and air quality. CV methods encompass personal interviews, mail surveys, and telephone surveys which elicit consumers' willingness to pay (WTP) for nonmarket goods "contingent" on a given hypothetical scenario. Carson *et al.*'s (1993) bibliography of CV studies provides a list of over fourteen-hundred studies. Therefore, collectively, CV researchers have a solid foundation for designing CV studies that are reliable (Randall 1987).

One important consideration when designing a CV survey is the elicitation method used. The relevant distinction is between continuous valuation methods including open-ended (OE) questions versus discrete methods such as iterative bidding and the dichotomous choice (DC) method. The payment card method can be considered as either a discrete method or a continuous method (Cameron and Huppert 1989). If the dollar values listed on the payment card are close together, the resulting data can be treated as essentially continuous.

Because CV estimates of WTP are based on survey responses, we must be concerned with how we ask the questions, and how respondents formulate their answers. Several studies have looked at whether continuous methods generate value estimates that are different from those generated by discrete methods. Hoehn and Randall (1987) showed that the DC method can provide larger values than the OE method because of differences in the value formation process. Additionally, Boyle *et al.*'s (1993) found that DC WTP estimates were greater than PC estimates. In contrast, Boyle and Bishop (1988) found

[1] This research was financially supported by the Economic Research Service, U.S. Department of Agriculture as part of a cooperative agreement (43-3AEK-2-80072) with the University of Kentucky (1992-93).

that the OE method provided larger values than the DC method. Several recent papers have further investigated these differences.

Kealy and Turner (1993) found that DC WTP estimates were 1.4 to 4 times larger than OE estimates for a public good but found no difference for a private good. They compared OE and DC methods using the same sample of respondents. They used a sophisticated estimation approach to account for correlation between the answers of these two questions due to variability in preferences. Still, their design could not test whether the presence of a DC question affected stated OE values, or vice versa. Their results did suggest that question order matters, though. They used a voluntary contribution payment vehicle for public good versus a price for private good. Contribution payment vehicles are not incentive compatible with expressing maximum willingness to pay. Overall, the results are interesting and suggestive.

Kriström (1993) showed that a large proportion of respondents said "YES" to a DC question with bid values that were larger than most stated OE values. Kriström states that understanding such disparities is important given the increasing popularity of the DC method in CV studies (p. 69). He introduced a simple way to compare continuous and discrete elicitation methods using a non-parametric test. The non-parametric approach avoids problems associated with choosing an inappropriate functional form.[2] He tested for cross contamination using one sample of respondents that answered both OE and DC questions and a sample that only received the OE question. His results suggest that, overall, the OE responses were not affected by the presence of the DC question. However, he did not test whether individual OE responses were impacted by the bid value posed in the DC question. Also, as with Kealy and Turner, he used a voluntary contribution payment vehicle.

It would be interesting if both the Kealy and Turner and the Kriström studies could be replicated using separate samples and incentive compatible payment vehicles. To avoid complications due to cross contamination of the OE and DC questions, it is preferable to apply the OE and DC

[2] A limitation of the non-parametric approach is that it does not allow consideration of the impact of demographics or other preference shifters on WTP.

methods to separate samples of respondents. Researchers do not know if using the same sample for two elicitation methods is a problem but it is cleaner to use separate samples to make comparisons.

Unlike Kealy and Turner and Kriström, Boyle *et al.* used separate samples for OE and DC questions. Boyle *et al.* did this in CV studies on the value of moose hunting and the value of oil spill protection. For three different valuation situations, they found that the mean WTP using the DC method was significantly higher than the mean WTP using the OE method in one case, significantly lower in one case, and found no significant difference in one case. Direct comparisons of the OE and DC means may be misleading, however. OE responses may not follow any standard distribution. Even if the question method does not impact stated values, so that the DC sample of respondents has the same distribution of values as the OE subsample, use of an inappropriate distribution when estimating the DC mean will generate biased estimates. A difference observed between the OE and DC means could be due to the statistical estimation of the mean from the DC data.

REASONS WHY WTP ESTIMATES MIGHT BE DIFFERENT

Two potential reasons why the mean WTP estimated from OE data are different from that estimated from DC data are differences in statistical method and differences in psychological behavior. By psychological behavior, we mean the mental processes respondents undergo when completing a CV question. Psychological behavior may influence WTP estimates in a variety of ways such as by negatively or positively influencing the mean WTP. Alwin and Krosnick (1991) concluded that both survey design and the respondents' personal characteristics affect the reliability of attitudinal survey measures. Several interesting papers have delved into the influence of the statistical method on mean WTP but few have contributed more than a cursory mention of the influence of psychological behavior.

Our paper hopes to accomplish two goals. The first objective is to observe to what extent differences in WTP estimates between the PC and DC elicitation methods are attributable to: (1) the respondents' psychological behavior, and (2) the statistical approach used. To this end, we used synthetic data techniques to isolate differences due to these components. We show that this procedure can be done correctly and that differences in WTP existing between elicitation methods are not largely due to differences in statistical approach but rather to psychological behavior. The second objective is to determine whether either the PC or the DC format is more sensitive to the risk level.

Two key papers have made strides towards understanding the influence of psychological behavior. Cameron and Huppert (1991) introduced an innovative approach whereby they used a sample of PC responses to generate "synthetic" DC data sets by randomly matching the PC responses to a fixed set of DC bid values. These synthetic data sets show how the PC respondents would have answered DC questions, assuming they retained their same WTP values. Cameron and Huppert used the synthetic DC data in lieu of actual dichotomous choice data for making comparisons between discrete and continuous elicitation methods. They showed that even if the distribution behind PC and synthetic DC data is identical, DC values can be quite large due to random sampling variability. Any difference between the mean calculated from the synthetic DC data set and the mean calculated directly from the PC data would be due to: (1) bias introduced by the statistical techniques used to analyze the DC data or (2) random effects due to sampling error. Cameron and Huppert found that DC mean WTP values can be much larger than the original mean WTP value, simply due to random variability, *i.e.*, luck of the draw. If we are to conclude from an observed difference between an PC mean and a DC mean that elicitation method impacts stated WTP, we must show that the difference is not due to random error or to bias in the analysis of the DC data. They did not emphasize the impact of psychological behavior on WTP estimates.

Boyle *et al.* generated synthetic DC data and extended the Cameron and Huppert study by comparing the synthetic DC data with actual DC data. In particular, they isolated differences in behavior between the OE and DC samples by using the same statistical approach to estimate mean WTP for the synthetic data sets and for their DC data sets. The synthetic data approach is better than a straight comparison between OE mean and DC mean, because OE values may not follow any standard distribution, in which case the statistical analysis is flawed. Using the synthetic approach on OE data introduces the same flaw for both the OE and DC samples.

Boyle *et al.* compared the synthetic DC data with the DC data and found that the DC WTP means were significantly higher than the synthetic DC values. They inferred that the two samples of respondents did demonstrate different stated values. Their analysis was complicated by the fact that their synthetic mean WTP estimates are quite a bit lower than their original OE mean WTP estimates. This result suggests that the assumed distribution used to analyze the discrete data sets was inappropriate, leading to biased WTP estimates. Boyle *et al.* assumed that WTP for the DC data sets were normally distributed across individuals. Yet, it is quite common to see skewed distributions of WTP. Kanninen and Kriström (1993) showed that there is a bias inherent in transforming the wrong bid structure. Their results would have been cleaner to interpret had they used a distributional form that more closely matched their OE data. These earlier papers are intriguing, yet there are many unresolved issues surrounding elicitation methods and the limited explanations about why WTP values differ across methods motivated this study.

In this study, we investigate the impact of elicitation method on estimated WTP to avoid residues of a specific pesticide on grapefruit. In our hypothetical scenario, the risk reduction associated with avoiding this pesticide is an excludable good, in that a consumer can avoid the pesticide only by purchasing a more expensive alternative. We generally follow the approach of Boyle *et al.* with slight differences in the statistical analyses, including use of an asymmetric distribution. We use separate samples for different elicitation methods and a different continuous question format (*i.e.*, PC) than the standard OE format. We show that the result holds for a private good, in contrast to Kealy and Turner.

REASONS WHY RESPONDENT BEHAVIOR MIGHT BE DIFFERENT

In an ideal CV survey situation, respondents would evaluate their preferences, formulate a value, and then answer either an OE or DC question consistently with that value. Differences between the OE and DC values show that this ideal has broken down somewhere. We identify four categories of breakdowns in the value formation/value statement process: (1) protest zeros/scenario rejection, (2) value understatement, (3) value overstatement, and (4) absolute yea-saying/scenario rejection.

With protest zeros/scenario rejection, respondents do not accept the hypothetical scenario.[3] This problem is widely recognized, and most researchers try to identify protest respondents in a DC context, and remove them from the sample. If protest respondents are removed from the DC sample, but not from the PC/OE sample, this would result in higher DC values. We leave them in for both samples just to be safe, and to make an even comparison.

Value understatement may occur due to conservatism. If respondents are unsure of their WTP in a PC or OE format, they may give lower WTP values to be on the safe side. Respondents may feel that it is worse to overbid than to underbid. If respondents do not want to make a false statement, they will have lied if they overbid. In contrast if the respondents' stated value lies below their true maximum WTP, they have not really lied in that they would still pay their stated value. To be safe, respondents may underbid to avoid lying. Switching to a DC format may do two things. First, the DC format may force more careful consideration of the WTP value when the bid is close to the true WTP. More careful consideration will reduce the variance in the subjective distribution for WTP, so that the conservative estimate ends up being higher. Second, the loss function may be more symmetric in the DC format giving less incentive for conservatism.

Value Overstatement may occur due to hypothetical nature of the payment. This argument has been made as "hypothetical dollars are cheaper than real dollars." Stated WTP in hypothetical dollars will therefore be larger than actual WTP. We see no reason why this effect would be different between OE (or PC) format and DC format.

Absolute yea-saying/scenario rejection could occur in either the DC or OE format. Some people may have difficulty understanding the trade-off that is being asked, or may have difficulty evaluating their own WTP. These people may just short circuit the decision problem and say "I do not know what its worth to me, all I know is that I want the amenity." Theoretically, such people would

[3] They believe they should not HAVE to pay, or that the money is not really needed. In a OE or PC context, they would put a value of $0.00 or refuse to answer the question. In a DC context, they say NO or do not return the survey. In our survey, the higher the bid, the less likely the respondent would complete the WTP question or return the survey. This implies that this breakdown may be especially prevalent at higher DC bids.

say yes to any bid asked. How would such people answer a PC or OE question? The PC or OE format forces direct consideration of the dollar amount. Respondents cannot simply ignore their own WTP values. This will force them to come up with some stated WTP value. Even if they do a poor job of value formation, their stated value will be finite. The DC format does not force stated WTP values to be finite. This implies that the respondent switches to a "no" value at a higher dollar value than what he would actually pay, *i.e.*, the proportion of the area in the right hand tail of the probability of a "YES" distribution is too high. Other researchers have made cursory references to yea-saying (Boyle *et al.*, Mitchell and Carson).

Of these four categories of breakdowns in the value formation/value statement process, two could result in DC values being greater than PC values: conservatism and yea-saying.[4] If it is caused more by conservatism, PC values will understate true values and DC values will be closer to true values. If it is caused more by yea-saying, DC values will overstate true values and PC values will be closer to true values. Either way, we cannot say whether or not the true value lies between the two estimates.

SURVEY DATA

This study uses data from 1671 returned contingent valuation mail surveys of U.S. grapefruit consumers (Buzby *et al.* 1993). The survey sample consisted of those grapefruit consumers contacted by phone who were willing to participate in a follow-up mail survey.[5] In the hypothetical market situation posed, the respondent faced a choice between buying a baseline grapefruit for fifty cents, or buying a relatively safer grapefruit at a higher price. Other than the pesticide residue risk, the two grapefruit were identical. The surveys incorporated risk ladders to help respondents understand the

---

[4] An unequal treatment of protests could also result in relatively larger DC estimates, but we correct for this.

[5] The phone survey used random digit dialing and Dillman's Total Design Method while the CV mail survey followed techniques outlined by Dillman and Mitchell and Carson.

relative risks. Loomis and Duvair (1992) found the risk ladder to be an effective tool for helping respondents answer contingent valuation questions involving risk changes.

Four survey versions paired two pesticide residue risk reductions (50% and 99+%) and the two elicitation methods (PC and DC). Most of the paper focuses on the 99+% risk reduction. The 50% risk reduction scenario was used to determine the robustness of the WTP estimates to a change in the risk scenario.

The payment card consisted of a column of values ranging from zero to fifty cents above the original purchase price of one baseline grapefruit.[6] We asked respondents to circle the one amount that shows the most that they would pay above the purchase price of a baseline grapefruit to buy one relatively safer grapefruit.

For the DC method, the price for the relatively safer grapefruit was randomly set at one of ten prices (bids).[7] We asked respondents which grapefruit they would buy given the prices stated and the differing risks.

The demographics of the survey sample tracked fairly well with the general population with the exception that there were more women and fewer children in the surveys. These differences were expected because the samples represent a population of "food shoppers" whereas the census represents the overall population. The response rate for the PC version was 79.5 % and 77.3 % for the DC version, suggesting that respondents were not put off by either format.[8]

---

[6] There were 31 values between zero and fifty cents on the PC column of values. Space was provided for respondents to fill in values above fifty cents.

[7] Ten prices for TBZ treated grapefruit were randomly used in the surveys: \$.53, \$.55, \$.60, \$.70, \$.80, \$1.00, \$1.25, \$1.50, \$2.00, and \$2.50. These prices were chosen after inspection of a sub-sample of PC responses. An equal number of surveys were assigned to each price.

[8] The response rates represent the percent of surveys returned out of the total number of surveys mailed. If undeliverable addresses were removed, response rates would increase to 79.5% and 77.3% respectively.

Payment Card Data

Prior to analysis, we applied a conservative two step procedure to the PC WTP data: (1) undeliverable surveys were dropped from the sample, and (2) all unit and item WTP nonresponses were assumed to have a zero WTP to correct for nonresponse bias. A simple mean and median were found for all PC WTP values. One interesting occurrence is that for the payment card method, WTP values tended to be centered around ten cent increments (*i.e.*, 10, 20, 30 40 . . .). This behavior may suggest that consumers round numbers for simplicity. To simulate the variability that we would see in the PC mean WTP, we constructed one-thousand bootstrapped data sets. Each of these was generated by sampling with replacement from the original data set of 667 PC WTP values.

Dichotomous Choice Data

For the dichotomous choice data set, we assume that each individual's WTP for the risk reduction is a random variable, with ln(WTP) following a logistic distribution. The probability that an individual respondent will say yes to a DC question given bid P is then

$$F(P) = \frac{1}{1 + e^{-(\alpha + \beta \ln(P))}}$$

This implies that WTP is a positive random variable with an asymmetric distribution, skewed to the left. Following Hanemman (1984), the mean WTP across individuals is found by integrating under F(P). We follow the conservative practice used by Bishop and Heberlein (1979) and Seller, Stoll, and Chavas (1985) of truncating the integration at the highest P used, $2.00, so that the calculated mean is given by

$$Mean\,WTP = \int_0^2 F(P)\,dP$$

As pointed out by Duffield and Patterson (1991), the effect of truncation is to set all WTP values greater than \$2.00 equal to \$2.00.

The parameters α and β were estimated using maximum likelihood logistic regression. Mean and median WTP were calculated from these maximum likelihood estimates. To obtain confidence intervals for both the mean and median WTP, we generated one thousand random draws from the estimated multinormal distribution of α and β, as described by Krinsky and Robb (1986).

Synthetic DC Data

To determine whether any differences between the means and medians calculated from the PC and the DC data are due to differences in the statistical approach used or differences in the respondents' psychological behavior, it is first necessary to construct a synthetic DC data set from the PC data that is identical in structure to the DC data set. To generate a synthetic DC data set, the bid from each observation from the DC data set was compared to a randomly chosen (with replacement) WTP value in the PC data set. For each DC observation, the respondent is assumed to say "yes" if the randomly chosen PC WTP value is greater than or equal to the DC bid, and "no" otherwise. Cameron and Huppert provide technical details of how this can be accomplished.

This process was repeated, generating one thousand synthetic data sets. For each data set, the parameters α and β were estimated using logistic regression. This process generated confidence intervals for the mean and median WTP. For exposition purposes in Figures 1 and 2, a "modal" synthetic DC data set characterizes the central tendency for these synthetic data sets. We generated the "modal" synthetic DC data set by calculating, for each bid level, the expected number of "yes"

STATISTICAL ANALYSIS

The synthetic data served two purposes. First, mean and median WTP estimated from the synthetic DC data were compared to estimates from the original PC data to determine whether the statistical method used biased the estimated WTP values when psychological behavior in answering the WTP question was held constant (*i.e.*, the same PC respondents). Second, estimates from the synthetic DC data were compared with estimates from the DC data to determine whether respondents exhibited different preferences (*i.e.*, when the DC statistical approach was held constant). Bootstrapped replications allow statistical testing of these differences. Before performing such tests, however, we performed a normality test suggested by Kmenta for the bootstrapped WTP estimates. The test results rejected the normality hypothesis for both PC and synthetic estimates. Thus, to test the differences of two empirical distributions, we must use more universally applicable tests than those based on normality.

Poe, Lossin, and Welsh (1993) have proposed a nonparametric type test based on the method of convolutions that does not require the assumptions of normality or approximate normality for the distributions. There are two difficulties in applying this test to the comparison of the WTP estimates from PC and DC data. First, the probability density functions of WTP are too complicated to be estimated from the empirical data. Second, the discrete approximations of convolution are very difficult to calculate using a software that does not have a convolution routine. Instead, we used a simpler approach that retains the spirit of their technique. We constructed a resampling-based test that can be easily conducted using bootstrap data. Let X^* and Y^* be vectors containing the one-thousand bootstrapped WTP estimates obtained from two different approaches. We randomly drew one thousand pairs of X and Y, and for each pair calculated X-Y. Our test statistic, p, is then the number of pairs that generate a negative difference divided by one-thousand. If p is less than α, then we reject the null hypothesis that the WTP estimates are equal, at a one-tailed significance level of α.

Table 1 provides measures of central tendency for the PC, DC, and synthetic DC data under the 99+% risk reduction scenario. The mean PC WTP was 19 cents per grapefruit above the starting price of fifty cents whereas the DC mean WTP was 69 cents per grapefruit above the starting price. Interestingly, the higher the DC bid, the less likely the respondent would continue to participate. The mean WTP for the synthetic data was 25 cents per grapefruit above the starting price, which lies between the PC and DC means. All three mean values were significantly different (p=0.028 for synthetic v. PC and P-0.002 for synthetic v. DC). Therefore, of the 50-cent difference between the PC mean and the DC mean, 6 cents was due to bias caused by the statistical analysis, and 44 cents was due to the differences in behavior. The confidence interval of the means for the DC data was broader than for the PC data. Medians were more stable across data types than the means, especially for the PC and synthetic DC medians. Median WTP from the synthetic data was significantly different from median WTP from the DC data (P<0.001) but not from the PC data.

Figure 1 illustrates these differences, showing the estimated logistic regression for the original DC data set and the regression for the modal synthetic data set. Figure 1 also shows the actual frequency of yes responses in the original DC set (as shown by the grey squares), and the cumulative distribution of the PC data. Figure 1 shows that the difference between the PC mean and the synthetic mean is due mostly to divergence in the range between 50 cents and $2.00, but that the DC regression diverges throughout the range of the data.

Robustness of WTP Estimates under a Different Risk Scenario

The third objective was to see if the above differences are robust to a change in the risk scenario by using the two CV versions under the lower risk scenario (50% reduction). Table 2 and Figure 2 contain the same information for the 50% risk reduction scenario. The same patterns were seen with both risk reduction levels. Here, Mean WTP was 16 cents for the PC data, 21 cents for the synthetic data, and 69 cents for the DC data. The synthetic mean was not significantly different for the PC mean (p=0.131) but was significantly different from the DC mean (p=0.020). Statistical comparisons of the medians generated the same results (p=0.794 for the synthetic V. PC and p=0.001 for synthetic V. DC).

To summarize, we found large differences between the estimated mean WTP from the DC data and from the PC data. Most of these differences were due to differences in how the respondents answered the questions. Only a small portion of the differences was due to bias from inappropriate assumptions in the statistical analysis, and the differences cannot be explained by "luck of the draw." Estimates of median WTP followed a similar pattern, but the differences across elicitation methods were smaller.

Finally, we used our significance test to compare WTP across the two risk reduction scenarios, to see whether respondents to either elicitation techniques were sensitive to the scenario presented. PC mean WTP was 3.5 cents higher for the 99+% scenario than for the 50% scenario, and this difference was statistically significant. The DC mean WTP estimates was only 0.7 cents higher for the 99+% scenario, and the difference was not statistically significant. Are DC respondents paying less attention to the scenario? Our inability to detect a difference may be because each DC response contains less information than each PC response, so that any statistical test will be less powerful. When we compare the synthetic means between the two risk reduction scenarios, we find that the 99+% scenario generates a mean WTP that is 4.5 cents higher than the 50% scenario, but that this difference is not statistically significant, although the original PC data did show a significant difference.

CONCLUDING REMARKS

While we used the payment card approach rather than open-ended questions to generate continuous data, and valued a private rather than a public good, our results were consistent with Hoehn and Randall and Boyle *et al.*'s results in that the discrete valuation method obtained higher WTP estimates than the continuous method. We could show that these differences cannot be explained by the "luck of the draw," and that most of the differences is due to differences in behavior between the two samples of respondents. Specifically, respondents were more likely to say "yes" to a given bid in the DC versions than in the PC versions. This effect was most pronounced at higher bid levels.

We feel that it is not enough to report means alone. Instead, reporting of both means and the whole WTP distribution is necessary to obtain a complete understanding of the differences between the PC and DC WTP data. The selection of the DC truncation point is critical because the lower the point, the less the PC and DC distributions will diverge. If researchers do not pay enough attention to the right hand tail of the WTP distribution, either statistical method could overestimate the mean.

Further research could expand knowledge about why respondent behavior might differ between question formats. Validity checks are needed to determine which method provides WTP values closest to the "true" values. Market simulation experiments that compare hypothetical responses to behavior involving actual exchange of money would shed insight as to the relative validity of the elicitation methods. More research is needed to asses the differences between elicitation methods. Yet, it is likely that neither the PC method nor the DC method is absolutely correct or absolutely wrong. The DC WTP values may be biased upwards by yea-saying and the PC WTP values could be biased downwards by ultra conservatism. Although WTP estimates were higher for the DC method than for the PC method, the true value could lie somewhere in between.

Table 1: Mean WTP, Median WTP, and 95% Confidence Intervals

Under the Higher Risk Reduction Scenario[a]

(in cents)

VERSION	MEAN WTP	MEDIAN WTP
Payment Card n= 667	19.2 (17.7, 20.8)	10 (10, 10)
Synthetic DC n=684	25.6 (19.4, 32.5)	(10.3) (6.9, 14.8)
Dichotomous Choice n=684	69.4 (61.8, 77.6)	21.5 (15.7, 29.2)

[a] The number in parenthesis is the confidence interval.

Table 2: Mean WTP, Median WTP, and 95% Confidence Intervals
Under the Lower Risk Reduction Scenario[a]

(in cents)

VERSION	MEAN WTP	MEDIAN WTP
Payment Card n=392	15.7 (14.0, 18.4)	10 (5, 10)
Synthetic DC n=388	21.1 (12.8, 30.5)	7.4 (3.8, 12.3)
Dichotomous Choice n=388	68.7 (58.9, 79.4)	17.0 (9.7, 26.6)

[a] The number in parenthesis is the confidence interval.

Figure 1. PC, Synthetic DC, and DC WTP Distributions for the 99+ Percent Risk Reduction Scenario

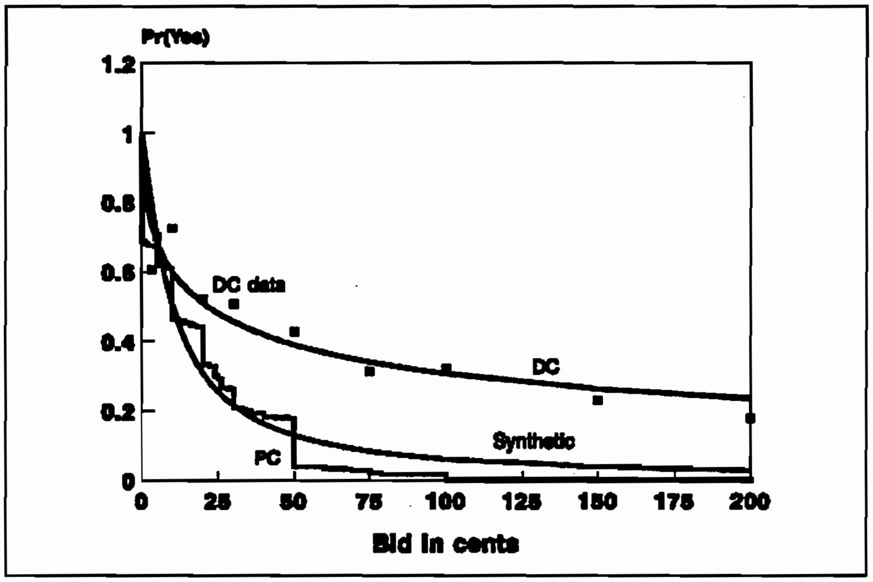
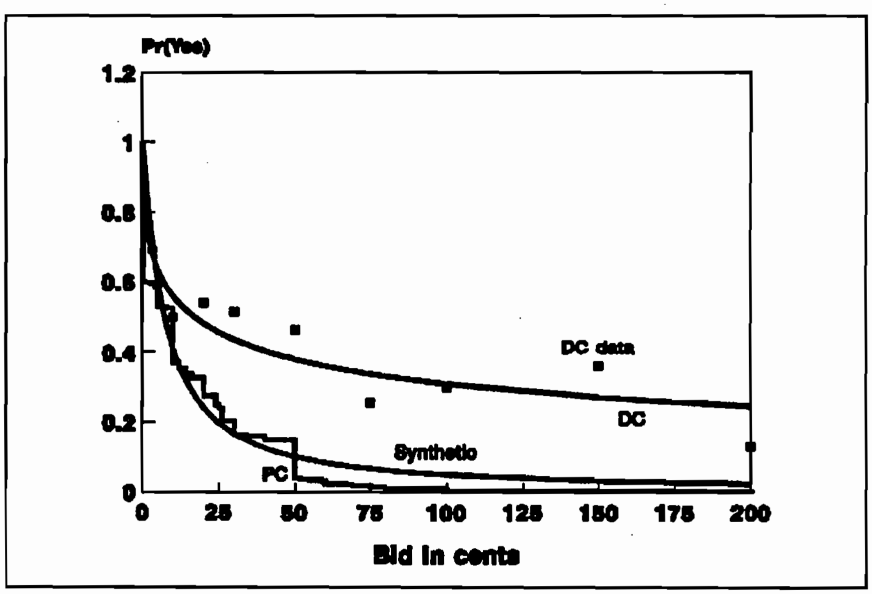


Figure 2. PC, Synthetic DC, and DC WTP Distributions for the 50 Percent Risk Reduction Scenario



REFERENCES

Alwin, Duane F., and Jon A. Krosnick. "The Reliability of Survey Attitude Measurement." *Sociological Methods and Research*, 20(1)(Aug. 1991):139-181.

Bishop, Richard C. and Thomas A. Heberlein. "Measuring Values of Extra-Market Goods: Are Indirect Methods Biased?" *Amer. J. of Agr. Econ.*, 61(Dec. 1979):926-930.

Boyle, Kevin J. and Richard C. Bishop. Welfare Measurements Using Contingent Valuation: A Comparison of Techniques. *Amer. J. of Agr. Econ.* 70(1988):20-28.

Boyle, Kevin J., F. Reed Johnson, Daniel W. McCollum, William H. Desvousges, Richard W. Dunford, and Sara P. Hudson. "Valuing Public Goods: Discrete Versus Continuous Contingent-Valuation Responses." Paper submitted for publication consideration in the *Amer. J. of Agr. Econ.*.

Buzby, Jean C., Jerry R. Skees, and Richard C. Ready "Using Contingent Valuation to Measure Consumers' Willingness to Pay for Food Safety: Results from a Survey on Pesticide Risks." Paper presented at the NE-165 workshop Valuing Food Safety and Nutrition, Old Town, VA, June 2-4, 1993.

Cameron, Trudy Ann, and Daniel Huppert. "OLS Versus ML Estimation on Non-market Resource Values with Payment Card Interval Data." *J. Envn. Econ. and Mngt.* 17(1989):230-246.

Cameron, Trudy Ann, and Daniel Huppert. "Referendum Contingent Valuation Estimates: Sensitivity to the Assignment of Offered Values." *J. of Amer. Stat. Assoc.* 86(1991):910:918.

Carson, R. T., N. Carson, A. Alberini, N. E. Flores, and J. Wright. *A Bibliography of Contingent Valuation Studies and Papers.* La Jolla, CA: Natural Resource Damage Assessment Inc., April 1993.

Dillman, Don A. *Mail and Telephone Surveys: The Total Design Method.* New York: John Wiley & Sons, 1978.

Duffeild, John W., and David A. Patterson. "Inference and Optimal Design for a Welfare Measure in Dichotomous Choice Contingent Valuation." *Land Econ.* 67,2(May 1991):225-39).

Hanemann, W. Michael, "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses," *Amer. J. of Agr. Econ.*, 66(1984):335-79.2

Hoehn, John P. and Alan Randall. "A Satisfactory Benefit Cost Indicator from Contingent Valuation." *J. of Envir. Econ. and Mngt.* 14(1987):226-247.

Kanninen, Barbara J. and Bengt Kriström. "Sensitivity of Willingness to Pay Estimates to Bid Design in Dichotomous Choice Valuation Models: Comment." Land Econ. 69,2(May 1993):199-202.

Kealy, Mary Jo, and Robert W. Turner. "A Test of the Equality of Closed-Ended and Open-Ended Contingent Valuations." *Amer. J. of Agr. Econ.*, 75(1993):321-331.

Kmenta, Jan. *Elements of Econometrics.* New York: Macmillan Pub. Co., Inc. New York, 1971.

Krinsky, I. and A. L. Robb. "On Approximating the Statistical Properties of Elasticities," Rev. of Econ. and Stat. 68(Nov. 1986):715-719.

Kriström, Bengt. "Comparing Continuous and Discrete Contingent Valuation Questions," Envir. and Res. Econ. 3,(1993):63-71.

Loomis, John, and Pierre Duvair. "Evaluating the Effect of Alternative Risk Communication Devices on Willingness to Pay: Results From a Dichotomous Choice Contingent Valuation Experiment," Land Econ. 69,3(1993):287-298.

Mitchell, Robert Cameron, and Richard T. Carson. *Using Surveys to Value Public Goods.* Washington, D.C.: Resources for the Future, 1989.

Poe, Gregory L., Eric K. Lossin, and Michael P. Welsh. "A Convolutions Approach to Measuring the Differences in Simulated Distributions: Application of Dichotomous Choice Contingent Valuation." Working Paper #93-03, Dept. Agr. Econ., New York State College of Agr. and Life Sciences.

Randall, Alan. *Resource Economics: An Economic Approach to Natural Resource and Environmental Policy.* New York: John Wiley & Son, 1987.

Sellar, C., J. R. Stoll and J. P. Chavas. "Validation of Empirical Measures of Welfare Change: A Comparison of Nonmarket Techniques." *Land Econ.* 61(1985):156-75.

van Ravenswaay, Eileen O. "Consumer Perception of Health Risks in Food." In *Increasing Understanding of Public Problems and Policies-1990.* Oak Brook, Ill.: Farm Foundation, 1990.

SOME TRADE-OFFS IN THE USE OF FOLLOW-UP QUESTIONING IN DICHOTOMOUS CHOICE VALUATION*

by

Joseph A. Herriges and Jason F. Shogren
Department of Economics
Iowa State University

Abstract

Follow-up questions have been proposed as one way of improving the efficiency of dichotomous choice contingent valuation questionnaires. In this paper, we investigate two broad issues regarding their application using data from a CV study of the Storm Lake Watershed in northcentral Iowa, which varied both the extent and form of follow-up questioning. First, we consider the impact of follow-up questioning on survey response rates and the ability of subjects to follow the complex questioning sequence involved. Second, we develop a model of starting point bias in the context of dichotomous choice questioning with follow-up. Monte Carlo simulation is used to illustrate the potential bias imparted to estimates of both the mean WTP and its dispersion in the population. We use the CV data to estimate the extent of starting point bias.

Preliminary Draft
Do not cite without Author's permission

April 8, 1994

Presented at the W-133 Regional Research Project
1994 Annual Meetings
Tucson, AZ

* This research was funded in part by the Leopold Center for Sustainable Agriculture, Iowa State University's Agricultural Experiment Station, and the W-133 Regional Research Project. The authors would like to thank participants at the W-133 meeting for their helpful comments, as well as Michael Flint, Terrance Hurley, Shagorika Talwar, and John Yeomans for their able research assistance.

SOME TRADE-OFFS IN THE USE OF FOLLOW-UP QUESTIONING IN DICHOTOMOUS CHOICE VALUATION

1. INTRODUCTION

Dichotomous choice (or closed-ended) question formats now dominate the contingent valuation (CV) of nonmarket goods and services. In their simplest form, survey respondents are asked if they are willing to pay (or willing to accept) a fixed sum of money for a specified change in the amenity of interest.[1] One advantage of this "take-it-or-leave-it" format is that it mimics the decision making task that individuals face in everyday market transactions. In contrast to open-ended question formats that ask for a maximum willingness-to-pay (WTP), dichotomous choice is typically viewed as being easier to respond to and avoids incentive compatibility problems inherent in open-ended questions.[2] The key disadvantage of the dichotomous choice format is that the resulting survey responses reveal little about an individual's willingness-to-pay. The simple "take-or-leave-it" question, also referred to as single-bounded dichotomous choice, provides only a single bound on each individual's WTP (i.e., whether it is above or below the dollar value in the question). As a result, relatively large survey sample sizes are required to precisely characterize a population's WTP, which may in turn make the cost of the CV study prohibitive.

Follow-up questions have been proposed as one way to improve the efficiency of dichotomous choice questionnaires.[3] Rather than asking whether an individual's WTP is simply above or below a single bid value, a sequence of questions is used to narrow the range of their true WTP. Figure 1 illustrates three follow-up formats. The most common format, "double-bounded dichotomous choice", first asks the individual if their WTP is above a bid value (say B_M). If the response is "yes", the follow-up question asks if their WTP is above a second and higher bid value (say $B_H > B_M$). If the response is "no", the follow-up

---

[1] Bishop and Heberlein (1979,1980) are credited with the development of the "take-it-or-leave-it" approach. Mitchell and Carson (1989, pp. 97-104) provide a review of the alternative elicitation formats used in the CV literature.

[2] See, for example, Mitchell and Carson (1989, p. 101) and Arrow *et. al.* (1993).

[3] Careful selection of the bid values used, and their variation among the survey participants, can also significantly enhance the efficiency of a CV study. See, for example, Cooper (1993) and Kanninen (1993a,b).

question asks if their WTP is above a lower bid value (say $B_L < B_M$).[4] This sequence of responses places the individual's WTP into one of four regions: $R_1 \equiv [0, B_L)$, $R_2 \equiv [B_L, B_M)$, $R_3 \equiv [B_M, B_H)$, or $R_4 \equiv [B_H, +\infty)$.[5] Hanemann, Loomis, and Kanninen (1991) demonstrate that, because of the additional information provided by the follow-up question, double-bounded dichotomous choice is asymptotically more efficient that the single-bounded alternative. Combs *et.al.* (1993) use a slightly different sequence of follow-up questions to achieve the same efficiency gains. In their "one-way-street up" format, illustrated in Figure 1, the first question asks the individual if their WTP exceeds the lowest bid value, B_L. If they answer "no", no further questions are asked. If they answer "yes", they are asked if their WTP exceeds the second bid value, B_M. Again, a "no" response stops the questioning, while a "yes" response leads to a third and final question, namely whether their WTP exceeds the highest bid value, B_H. This one-way street format yields the same information as the double-bounded approach, placing the individual's WTP into one of the four regions defined above. Only the sequence of questioning has changed.

This paper investigates two potential costs of follow-up questioning, reduced response rates and starting point bias, which must be traded-off against their beneficial impact on survey efficiency. The efficiency gains attributed to follow-up questions are due to the additional information they provide, narrowing the uncertainty about each individual's WTP. These gains may be offset, however, if the use of follow-up questions reduces the response rates to the survey or to specific (WTP) questions within the survey. Second, many of the studies employing follow-up questions (Hanemann, Loomis, and Kanninen, 1991; McFadden and Leonard, 1993) have noted significant differences between the responses to the first question and the responses to follow-up questions. While a number of hypotheses have been put forward to explain this finding, we focus on the potential impact of starting point bias.[6] As Hanemann, Loomis, and

[4] This follow-up format was first proposed by Hanemann (1985) and Carson (1985) and first implemented by Carson, Hanemann, and Mitchell (1986). Subsequent applications include Cameron and James (1987) and McFadden and Leonard (1993).
[5] We assume that the individual's true willingness to pay is bounded from below by 0. WTP may also be bounded from above by individual's income or other financial constraints.
[6] Hanemann, Loomis, and Kanninen (1991, p 1262), for example, suggest the possibilities of yea-saying, weariness, and respondent uncertainty regarding the exact value of their WTP. More recently, Alberini,

Kanninen (1991, p. 1253) note, the use of follow-up questions represents a compromise between single-bounded methods and traditional bidding approach. Since the bidding method has long been criticized as being subject to starting point bias, it would seem that the limited form of bidding used in follow-up questions may also lead to biased responses.[7]

The remainder of the paper is divided into four sections. Section 2 describes a CV study of water quality in Storm Lake, Iowa, a popular recreational fishing site in the northcentral portion of the state. The study incorporates an experiment specifically designed to address the issues identified above. The survey sample is divided between households receiving dichotomous choice questionnaires with and without follow-up, with three variations in the format of the follow-up questions. Data from the study are then used in Section 3 to test the impact of the follow-up questions on the pattern of survey responses. Section 4 focuses on starting point bias in dichotomous choice questionnaires using follow-up. A model of starting point bias is developed and estimated using data from the Storm Lake study. The paper is concluded in Section 5.

2. EXPERIMENTAL DESIGN

The CV survey used in this paper to analyze follow-up questions was developed as part of a broader interdisciplinary effort to design and evaluate alternative riparian management practices. Specifically, the project focuses on the use of riparian buffer strips to reduce sediment flows within the Storm Lake watershed.[8] Storm Lake is a shallow glacial lake in northcentral Iowa that serves as a popular fishing and boating site in the region. While covering an area of approximately 3,000 acres, the lake is only eight feet deep on average and suffers one of the highest levels of turbidity in the state.[9] Despite earlier dredging efforts, fishing and boating activities are threatened by continued sediment flows from industrial,

---

Kanninen, and Carson (1994) and Cameron and Quiggin (1993) have suggested that transitory elements may exist exist in an individual's WTP and proposed the use of bivariate probit to control for these effects.

[7] See Cummings, Brookshire, and Schulze (1986) and Mitchell and Carson (1989) for summaries of the starting point bias literature.

[8] Menzel and Schulz (1993) provide a more detailed overview of the Storm Lake Project and the riparian management practices being considered.

[9] Jones and Bachmann (1978).

agricultural, and residential sources. Contingent valuation is being used to value the potential water quality improvements that would stem from installing riparian buffer strips and to measure the willingness of farmers to adopt the required production practices. Towards these ends, surveys have been administered to three target populations: (1) local residents, (2) recreationists, and (3) farmers within the watershed. For the purposes of this paper, we restrict our attention to the local resident and recreationist surveys.

Both the local resident and recreationist surveys were administered by mail.[10] Local residents were selected at random from the local telephone directory, while recreationists were initially contacted in person at the lake and later mailed the full survey instrument.[11] In order to encourage a high response rate, participants were promised a four dollar check when we received their completed survey. In addition, three follow-up mailings were used, including a postcard reminder one week after the initial survey mailing and new copies of the survey mailed two weeks and six weeks after the initial contact. Local resident surveys were conducted in late 1992, while the recreationists were mailed surveys in late 1993.

The survey instruments themselves were structured around three sections. The first section asked respondents to characterize their use of Storm Lake and other recreational sites within the region. Section two described the current physical characteristics of Storm Lake (including lake depth, turbidity, and catch rates) and its projected deterioration over the next decade. Subjects were then asked to consider three programs that would (1) maintain the current lake quality, (2) improve upon the current lake quality, and (3) postpone the lakes deterioration path. Dichotomous choice CV questions were used to bound each respondent's WTP for the proposed programs. For example, in the case of the first program, subjects were asked: "Would you be willing to pay $____ on a one time basis (payable in installments of $____ over the

[10] Typically, follow-up questions have been used in telephone surveys, where the flow of questioning can be controlled by the surveyor. Their use in a mail survey application raises additional questions about the ability and willingness of respondents to follow the more complicated survey instructions. We consider these issues in section 3 of the paper.

[11] Recreationists were contacted at Storm Lake through a series of sixteen eight-hour visits to the lake. The selection and timing of the visits were designed to correspond to the Iowa Department of Natural Resource's annual creel survey of the lake. The visits systematically varied the time-of-day, day-of-the-week, and season during which visitors were contacted. Recreationists were briefly asked about their current recreation activity and to provide their name and address for further surveying. Roughly 1100 contacts were made, with over 1000 individuals agreeing to provide the requested information.

next five years) to...maintain Storm Lake's current condition?" Depending upon their response to this question and the form of follow-up question being used, the same basic wording was use to narrow the range of their WTP.[12] The third section of the survey was used to elicit socio-economic information, including household income, education, family size, and employment status. A focus group was used to check clarity of the survey instrument.

The unique feature of the Storm Lake Project's CV effort lies in the controlled variation in the follow-up question formats. Table 1 illustrates the basic experimental design. Table 2 lists the bid values used for the program to maintain the current water quality and the variation of the bids used in the sample.[13] All 300 local resident surveys employed follow-up questions to elicit WTP for water quality improvements, but varied the format of the follow-up questioning. The local resident subjects were randomly assigned to one of three follow-up formats: "one-way up", "double-bounded", and "one-way down". This same division was used for half of the recreationist surveys. In addition, 300 of the recreationist surveys elicited WTP values *without* follow-up. This design allows us to address two broad issues - the response pattern impacts of follow-up questioning and starting point bias. First, by comparing the recreationist surveys with and without follow-up, we can investigate the impact of the additional follow-up questions (and their format) on the pattern of survey responses. Second, both the recreationist and local resident samples can be used to compare responses to the first dichotomous choice question to the follow-up question responses, testing for starting point bias and its dependence on the follow-up format.

3. RESPONSE PATTERNS

The benefit of follow-up questions lies in the increased information they provide about each individual's willingness-to-pay. These gains may be reduced if the additional length and complexity of the survey instrument either reduces the overall survey response rate or the response rate to individual WTP

---

12 See Appendix A for the exact wording of the WTP questions used for the three follow-up formats.
13 Different sets of bid values were used for the other two program, with higher bids for the lake improvement option and lower bid values for the deterioration postponement option. The bid values used in the recreationist survey were designed to mimic the distribution of responses found in the local resident survey returns.

questions. In the case of the Storm Lake recreationists, the survey instruments were twelve pages when follow-up questions were used and eleven pages when they were excluded.[14] With the exception of the follow-up questions themselves, the survey instruments were otherwise identical.

Figure 2 illustrates cumulative response rates with and without follow-up at each stage of the survey administration process. For example, the 59 percent response rate for the follow-up category and above "postcard reminder" in Figure 2 represents the total response rate from the time of the initial survey mailing but *prior to* the second mailing. This response is broadly attributed to the combination of the first mailing and the postcard reminder. The response rates were uniformly smaller when follow-up questions were included in the survey instrument - 81 percent versus 86 percent by the end of the survey process. With the exception of the postcard reminder period, the departures from the "without follow-up" form of the survey were on the order of six percentage points and statistically significant at a 5 percent significance level. Clearly, these reduced response rates diminish, and can overwhelm, the efficiency gains from the inclusion of follow-up question and increase the potential for nonresponse bias in the survey data. Additional research is needed to test the robustness of these response rate reductions and to establish the trade-offs between the information gains and response rate losses of follow-up questioning.

An second concern in the case of a mail survey application of follow-up questions stems from the complexity of the survey instructions. The standard telephone application allows the interviewer to walk the subject through the proper sequence of questions, skipping questions that are irrelevant. This is not the case for mail surveys. Although our focus group indicated that subjects were comfortable with the instructions sequencing them through the follow-up questions, results from both the recreationists and local residents surveys indicate that individuals did not uniformly respond to the questions as they were asked to. To illustrate this, we divide survey respondents into four categories. The "complete" category represents respondents who answered all of the questions they were supposed to answer, no more and no less.[15] The

[14] A page is defined as one side of an 8.5 inch by 5.5 inch sheet of paper.

[15] As one participant at the W-133 meeting noted, one cannot tell in the case of a mail survey whether the individual answered the questions in the desired sequence, but only whether their answers are consistent with the desired sequence.

"incomplete" category represents respondents who failed to complete the sequence of follow-up questions, but provide some information regarding their WTP. The final two categoried cover customers who answered more questions than they were supposed to. For example, a household in the "one-way street up" category who answered "no" to the lowest bid value should have stopped, but some subjects continued on to the middle bid value. If a household answered more questions then they were supposed to, but did so in a consistent fashion (e.g., "no" to B_L, "no" to B_M, and "no" to B_H) they are grouped in the "overkill" category. If, however, their responses were inconsistent (e.g., "no" to B_L, "yes" to B_M, and "no" to B_H), then they are categorize in "inconsistent" category.

Figure 3a provides a summary of these four response categories for the recreationist survey.[16] The good news is that few of the respondents fell into either the incomplete or inconsistent categories. The incomplete responses would reduce the efficiency gains from the use of follow-up questions, while the inconsistent responses would bring into question the responses themselves. The bad news is the high frequency of "overkill" responses in the "one-way up" and "double-bounded" formats, which typically took the form of three "no" responses (i.e., "no" to B_L, "no" to B_M, and "no" to B_H). One explanation for this result is that these represent protest or nay-saying responses, and not the subject's true WTP. However, the results in Figure 3b suggest that this is not the case - excluding protest bids does not alter this pattern of responses. The WTP questions in the survey were followed with a question asking the individual to provide a reason for their responses. Figure 3b excludes subjects who indicated that the plans were "unrealistic" or should be paid for by other funding sources. The pattern of responses does not change significantly once these "protest" responses are excluded from the sample. Further research is needed to determine the cause the "overkill" response.

4. STARTING POINT BIAS

As Hanemann, Loomis, and Kanninen (1991) note, the double-bounded dichotomous choice format represents a limited and more structured version of the traditional bidding approach to eliciting nonmarket

[16] A similar pattern was found for the local residents.

valuations. This same characterization applies to the "one-way street" follow-up formats. Yet the bidding game approach has long been criticized as subject to starting point bias (e.g., Mitchell and Carson, 1985, 1989; Cummings, Brookshire, and Schulze, 1986; Boyle, Bishop, and Welsh, 1985). In particular, the concern is that the initial bid provides a focal point or anchor for the uncertain respondent. "Confronted with a dollar figure in a situation where he is uncertain about an amenity's value, a respondent may regard the proposed amount as conveying an approximate value of the amenity's true value and anchor his WTP amount on the proposed amount." (Mitchell and Carson, 1989, p. 24). In this section, we first develop a model of this anchoring effect in context of dichotomous choice questionnaires with follow-up. Monte Carlo simulation is then used to illustrate the potential bias anchoring imparts on the estimated mean and standard deviation characterizing a population's WTP distribution. Finally, The local resident and recreationist survey responses are used to estimate the extent of anchoring in the Storm Lake Project's CV analysis.

4.1 Anchoring Effects: Theory and Simulation

Let W denote an individual true willingness to pay for a given water quality improvement program. The W's are assumed to be random draws from an underlying population distribution, with

$$W(X) = f(X;\theta,\varepsilon) \tag{1}$$

where X is a vector of individual characteristics affecting their WTP (e.g., income, education, age, etc.), θ is a vector unknown parameters, and ε is a zero mean random variate. [17] In the case of dichotomous choice CV, the W's are not directly observed. Instead, survey respondents reveal boundaries on their WTP's. Consider the first question in a dichotomous choice question using follow-up. When asked if they are willing to pay b_1 for a proposed water quality improvement, the subject's response reveals the value of the indicator variable, $d_1(b_1)$, where[18]

[17] In general, the W's will also depend upon characteristics of the water quality program being offered. Since, in the remainder of this paper we will focus on only the program for maintaining current lake conditions, these characteristics are suppressed for the sake of notational simplicity

[18] We use b_j to denote the j^{th} bid value in a sequence of discrete choice questions, as distinguished from B_L, B_M, and B_H, the low, medium, and high bid values used in various sequence combinations by the one-way street and double-bounded follow-up formats. Thus, for the one-way street up format, $b_1 = B_L$, $b_2 = B_M$, and $b_3 = B_H$, whereas for the one-way street down $b_1 = B_H$, $b_2 = B_M$, and $b_3 = B_L$.

$$d_1(b_1) \equiv \begin{cases} 1 & W \geq b_1 \\ 0 & W < b_1 \end{cases} \qquad (2)$$

Figure 4a illustrates the implications of the survey response in terms of positioning W within the underlying population distribution, with a "yes" response (i.e., $d_1 = 1$) placing it in the upper portion of the distribution and a "no" response (i.e., $d_1 = 0$) placing it in the lower portion.

The purpose of follow-up questioning is to further narrow the boundaries on W. For example, with $d_1 = 0$, a follow-up question might ask the individual if their WTP exceeds $b_2 < b_1$. The standard procedure is to treat this response in the same manner as the first discrete choice question; i.e., as revealing the value of an indicator variable $d_2(b_2)$, where

$$d_2(b_2) \equiv \begin{cases} 1 & W \geq b_2 \\ 0 & W < b_2 \end{cases} \qquad (3)$$

Thus, a "no" to the initial dichotomous choice question ($d_1 = 0$) together with a "yes" to the follow-up question ($d_2 = 1$) would restrict W to the shaded region in Figure 4b.

The problem with this approach is that it ignores the potential anchoring effect of the first bid offer on the subject's response to the follow-up question. As suggested in the literature on bidding games, an uncertain subject may view the starting bid (b_1) as providing information on the "correct" WTP value. Rather than comparing their true WTP to the follow-up bid, the respondent combines their true WTP with this new information (b_1) to form a revised WTP ($\tilde{W}$) and compares $\tilde{W}$ to b_2 in answering the following question. Formally, we model $\tilde{W}$ as a weighted average of W and b_1, with[19]

$$\tilde{W} \equiv (1-\gamma)W + \gamma b_1, \qquad (4)$$

where γ measures the extent of anchoring and $0 \leq \gamma \leq 1$. At extremes, $\gamma = 0$ implies no anchoring effect with $\tilde{W}$=W and $\gamma = 1$ implies that the subject ignores their prior WTP, replacing it with the $\tilde{W}$=b_1.

[19] This linear formulation of the anchoring effect represents but one way in which the respondent might combine the first bid with their true WTP. In section 4 below, a geometric weighted average is also illustrated.

Accounting for the anchoring effect, the follow-up question no longer provides information on the indicator variable d_2, but instead reveals the value of

$$\begin{aligned}\tilde{d}_2(b_2) &\equiv \begin{cases}1 & \tilde{W} \geq b_2 \\ 0 & \tilde{W} < b_2\end{cases} \\ &= \begin{cases}1 & (1-\gamma)W + \gamma b_1 \geq b_2 \\ 0 & (1-\gamma)W + \gamma b_1 < b_2\end{cases}\end{aligned} \tag{5}$$

Rewriting the inequalities of equation (5) in terms of W yields

$$\tilde{d}_2(b_2) = \begin{cases}1 & W \geq \tilde{b}_2 \\ 0 & W < \tilde{b}_2\end{cases} \tag{6}$$

where[20]

$$\tilde{b}_2 \equiv \frac{b_2 - \gamma b_1}{1-\gamma} \tag{7}$$

Comparing equations (3) and (6), it is clear that

$$\tilde{d}_2(b_2) = d_2(\tilde{b}_2). \tag{8}$$

That is, the net impact of anchoring on the information revealed by follow-up questioning is to replace the design bids (b_2) with what we refer to as the "effective" bid values $(\tilde{b}_2)$.[21] Furthermore, the direction of the movement in the bids is clear, since

$$\left|\tilde{b}_2 - b_1\right| = \frac{|b_2 - b_1|}{1-\gamma}. \tag{9}$$

Anchoring effectively widens the boundaries placed on W by the follow-up questions. The greater the anchoring effect (γ), the wider these boundaries become and, hence, the less information provided by follow-up questioning. This effect can be seen visually in Figure 4c, where again we use a follow-up question with a design bid $b_2 < b_1$. A "yes" response to the follow-up question would place the individual's

---

[20] This assumes that $\gamma < 1$.

[21] That is, a follow-up question using bid values of b_2 when anchoring occurs reveals the same information as a follow-up question using bid values of $\tilde{b}_2$ when anchoring does not occur.

W in the lightly shaded region of Figure 4c (i.e., $W \in [\tilde{b}_2, b_1)$), a much broader region than the one designed into the question sequence (i.e., $[b_2, b_1)$).

There are two implications of starting point bias in terms of the use of follow-up questioning. First, if starting point bias is ignored, estimates of mean WTP and its dispersion in the population will be potentially biased. The estimated mean WTP will be biased since the starting bid will draw the survey responses away from their true WTP values and towards b_1. The estimated dispersion of WTP in the population will also be biased by the analyst erroneously "squeezing" the revealed WTP's between the relatively narrow design boundaries (i.e., the b_i's), when the surveys actually place individuals within the "effective" boundaries (i.e., the $\tilde{b}_i$'s). For example, in Figure 4c, an analyst ignoring the anchoring effect would erroneously squeeze observations in the shaded region between the design boundaries of b_2 and b_1, in this case biasing the dispersion estimates downwards.

In order to better understand the extent to which anchoring can bias estimates of both the mean WTP and its dispersion in the population, a simple simulation exercises was conducted, varying the degree of anchoring and the follow-up format. The simulations were based upon the following model specifications:

- $W = f(X;\theta,\varepsilon) = \alpha + \varepsilon$, with $\alpha = 200$, $\varepsilon \sim N(0,\sigma)$, and $\sigma = 50$.
- $B_L = 100$, $B_M = 200$, and $B_H = 300$.
- N (the survey sample size) = 5000.

For each combination of follow-up design ("one-way up", "double-bounded", and "one-way down") and each level of anchoring ($\gamma = 0$ to .5 in steps of .025), 100 samples of size N were drawn from a N(200,50) distribution. For each these sample, the effective bids ($\tilde{b}_i$'s) were formed using equation (7). These in turn were used to determine the individual observation's "response" to the sequence of survey questions (i.e., $\tilde{d}_i$'s of equation 6). However, the design (b_i's) and *not* the effective bids were then used to estimate the mean and dispersion of the population's WTP. This mimics the behavior of an analyst ignoring starting point bias. Figure 5 shows the average over the 100 samples for each format and anchoring level combination.

As Figure 5a illustrates, both of the one-way street formats yield biased estimates of the mean WTP (i.e., $\hat{\alpha}$). In the one-way street up format the mean WTP is drawn further down towards the initial bid (B_L = 100) as the anchoring effect increases, with the mean WTP estimated to be roughly 150 (rather than its true value of 200) when $\gamma = .5$. A similar pattern emerges for the one-way street down format. In the case of the double-bounded approach, no bias arises. This is due to the fact that the initial bid (B_M) is equal to the true mean WTP and the population is distributed symmetrically around this mean.[22] The starting point bias squeezes the distribution tightly around the mean, but does not bias the estimated mean WTP. This is a potential benefit of the double-bounded format, but one that relies on both a well chosen starting value and a symmetric WTP distribution.

Figure 5b illustrates the impact that starting point bias can have the estimated dispersion of WTP in the target population (i.e., $\hat{\sigma}$). The effect is the same for all three formats, with the estimated dispersion coefficient being further biased downward as the anchoring effect increases. This bias can have potentially significant policy implications by suggesting that there is greater uniformity in WTP among a target population than is actually present.

Finally, we note that, even if the analyst controls for starting point bias, the efficiency gains from follow-up questioning is likely to be diminished by anchoring. The effect of anchoring is to spread the boundaries revealed by follow-up questioning. Bids that are optimal at the design stage are shifted to their "effective" bid levels by anchoring and become suboptimal.[23] However, it is important to emphasize that follow-up questions still provide efficiency gains over their single-bounded counterpart since they provide additional information without altering the information provided by the initial dichotomous choice question.

---

[22] In similar simulation exercises with $B_M \neq 200$ the double-bounded format was also found to be yield biased estimates of the mean WTP.

[23] It is of course possible that anchoring could make a suboptimal design more efficient by effectively shifting the bid boundaries revealed in the survey or that an optimal design with anchoring could be developed controlling for a specific level of bids.

4.2 Anchoring Estimates from the Field

To determine if anchoring is in fact an issue in applied research, we now estimate the extent of anchoring effects in the Storm Lake Project's local resident and recreationist CV surveys. Three models are estimated for each data base. First, a traditional single-bounded dichotomous choice model is estimated using responses to the first bid offer in each survey. Second, the information contained in the follow-up questions is used to estimate a multiple-bounded dichotomous choice model of WTP ignoring the potential for anchoring. Third, the extent of anchoring (γ) is measured directly by substituting the "effective" bid function of equation (7) in for the design bids in the multiple-bounded dichotomous choice model.

As noted in the literature, estimates of WTP can be sensitive to distributional assumptions. In our analysis to date, we have restricted our attention to the normal and lognormal distributions. The lognormal specification appears to better fit the skewed pattern of survey responses. In this case, it is convenient to specify the error term ε as entering the WTP equation (1) multiplicatively; i.e.,

$$W = f(X;\theta,\varepsilon) = g(X;\theta)\varepsilon \tag{10}$$

In addition, the linear anchoring process in equation (4) is replaced with a logarithmic specification, with

$$\ln\widetilde{W} = (1-\gamma)\ln W + \gamma \ln b_1 \tag{11}$$

Following the same steps as in the previous subsection, but using equation (11) in lieu of equation (4), it is straightforward to show that the k^{th} dichotomous choice question reveals the value of the discrete choice variable, $\widetilde{d}_k(b_k)$, where

$$\widetilde{d}_k(b_k) = \begin{cases} 1 & \ln W \geq \widetilde{\beta}_k \\ 0 & \ln W < \widetilde{\beta}_k \end{cases} \tag{12}$$

with

$$\widetilde{\beta}_k \equiv \frac{\ln b_k - \gamma \ln b_1}{(1-\gamma)} \tag{13}$$

The parameter $\widetilde{\beta}_k$ is the logarithmic version of the effective bid boundary $\widetilde{b}_k$. Finally, to focus attention on the anchoring effect itself, we restrict the nonstochastic portion of the WTP distribution to a constant, with:

$$g(X;\theta) = \exp(\alpha) \tag{14}$$

Using equations (10), (13) and (14), equation (12) can be rewritten as:

$$\tilde{d}_k(b_k) = \begin{cases} 1 & \ln\varepsilon \geq \dfrac{\ln b_k - \gamma \ln b_1}{(1-\gamma)} - \alpha \\ \\ 0 & \ln\varepsilon < \dfrac{\ln b_k - \gamma \ln b_1}{(1-\gamma)} - \alpha \end{cases} \tag{15}$$

where $\ln(\varepsilon) \sim N(0,\sigma)$. The resulting log-likelihood functions for the dichotomous choice models with and without follow-up are provided in Appendix B.

Table 3 provides maximum likelihoods estimates of α, σ, and γ for both the local resident and recreationist survey data bases and for three model specifications. The first column under each survey group uses data from the first bid response only to estimate parameters of the model. This corresponds to the traditional single-bounded approach. The second column provides estimates using the follow-up bid responses, but constraining the anchoring effect to be zero (i.e., $\gamma = 0$). This corresponds to the standard treatment of follow-up question. Finally, column three relaxes the constraint on γ, allowing the extent of anchoring to be estimated within the model.

Focussing first on the local residents, there appears to be little evidence of anchoring. The median WTP for maintaining the current level of water quality in Storm Lake is estimated to be approximately $150 using only the first bid responses. This point estimate does not change dramatically with the inclusion of the follow-up responses in column two, dropping the median WTP to just under $140. We do see the expected improvement in the precision of our parameter estimates. However, as the result of the previous subsection demonstrate, the stability of the median WTP estimate is not, in itself, proof that anchoring has not occurred. This is particularly true if the initial bids are centered around the true median WTP, as appears to be the case with local residents. We would, however, expect anchoring to reduce the estimated dispersion coefficient (σ) and this is not the case. The estimate of σ is virtually unchanged between columns one and two, again suggesting that anchoring is not a significant problem. Finally, when we estimate the anchoring parameter directly in column three of Table 3, it is found to be indistinguishable from zero at any reasonable significance level.

Turning to the recreationist survey (columns four through six of Table 3) a different picture emerges. Anchoring appears to play a significant role in these survey responses. First, we see a substantial shift in the estimated median WTP when the follow-up responses are used in the estimation and anchoring is ignored (i.e., $\gamma = 0$). Relying only on the initial bids (column four), the median WTP is approximately $42 for recreationists. Once the follow-up responses are included (column five), the estimated median value is increased by fifty percent to almost $65. This is consistent with hypothesis that survey respondents are anchoring to the higher initial bid values used in the survey (see Table 2). The dispersion coefficient estimates also suggest the presence of anchoring, with $\hat{\sigma}$ dropping by 37 percent (from 2.35 to 1.48) when the follow-up responses are included in the estimation. Finally, when the anchoring coefficient γ is estimated directly (column six), it is statistically significant at a 1 percent level, with $\hat{\gamma} = .36$. By correcting for the anchoring effect, our point estimates of the median WTP and the dispersion coefficient (σ) are again consistent with the first bid estimates. Including the follow-up questions still improves the precision of these estimates, but the extent of the improvement has been reduced once the anchoring effect is controlled for. When anchoring is not controlled for, the median WTP is seriously overstated, potentially leading to erroneously policy conclusion regarding the clean-up of Storm Lake.

5. CONCLUSIONS

Follow-up questions provide one mechanism to improve the efficiency of the dichotomous choice questionaire format used in contingent valuation. However, the results from our Storm Lake analysis suggest that there are also costs associated with their use. First, the additional complexity of the questionnaire may discourage survey response, directly reducing the efficiency gains from follow-up questioning and increasing the potential for nonresponse bias. The Storm Lake study found a five to seven percentage point reduction in the recreationist response rate when follow-up questions were included in the survey. Second, since follow-up questioning represents a limited form of the traditional bidding game approach to valuation elicitation, it may also be subject to the starting point bias found in earlier bidding game studies. The model presented in this paper suggests that if the individual does anchor their WTP to the initial bid in a sequence of bids, both the estimated median WTP and the estimated dispersion of WTP in the population can be significantly biased. Even if the analyst corrects for this anchoring effect, the efficiency gains from follow-up questioning is likely to be reduced, since the effective infomation content of the follow-up questioning is dilluted by the anchoring phenomenon. In the Storm Lake Project, anchoring was not found to be a significant problem for local residents. However, it did significantly bias both the estimated recreationist's WTP and the estimate of the dispersion of WTP among recreationists.

There are a number of directions for future research in this area. First, the preliminary analysis presented in this paper does not evaluate the impact that the follow-up questioning format has on the extent of anchoring. In addition, our model of the anchoring process assumes that it is constant across all individuals. The effect may depend upon the sequence of questioning and on the characteristics of the survey respondents (e.g., education, familiarity with the good being valued, etc.). Second, additional simulation work is needed to investigate the role of the underlying WTP distribution and bid design on the impact of starting point bias.

APPENDIX A: WTP QUESTIONS

In this appendix, we list the WTP questions used for the program to maintain Storm Lake's current water quality (Plan A) for each format. The dichotomous choice questions varied among the survey formats only in terms of the instructions used to step the participant through the sequence of follow-up questions. The bid variables (B_L, B_M, and B_H) and their annual counterparts ($A_i \equiv B_i/5$) are listed in italics below, but were replaced by their design values in the actual surveys.

A.1 One-Way Up

1. If the current rates of drainage and sedimentation continue throughout the next decade, the water quality of Storm Lake will change to:

Fish catch per hour (May/June):	**1 fish every 4 hours**
Average lake depth:	**7 feet**
Muck on lake bottom:	**20 inches**
Water clarity :	**Objects distinguishable ½ foot under water**

a. Would you be willing to pay **$$B_L$** on a one time basis (payable in installments of $$A_L$ annually over the next five years) to avoid these changes and maintain Storm Lakes' current condition?

❑ NO ➔ If no, please move on to Plan B

❑ YES➔ If yes, please continue to part b of this question

b. Would you be willing to pay **$$B_M$** on a one time basis (payable in installments of $$A_M$ annually over the next five years) to avoid these changes?

❑ NO ➔ If no, please move on to Plan B

❑ YES➔ If yes, please continue to part c of this question

c. Would you be willing to pay **$$B_H$** on a one time basis (payable in installments of $$A_H$ annually over the next five years) to avoid these changes?

❑ NO

❑ YES

A.2 One-Way Down

1. If the current rates of drainage and sedimentation continue throughout the next decade, the water quality of Storm Lake will change to:

Fish catch per hour (May/June):	**1 fish every 4 hours**
Average lake depth:	**7 feet**
Muck on lake bottom:	**20 inches**
Water clarity :	**Objects distinguishable ½ foot under water**

a. Would you be willing to pay $\boldsymbol{\$B_H}$ on a one time basis (payable in installments of $\$A_H$ annually over the next five years) to avoid these changes and maintain Storm Lakes' current condition?

- ❑ NO ➔ If no, please continue to part b of this question
- ❑ YES➔ If yes, please move on to Plan B

b. Would you be willing to pay $\boldsymbol{\$B_M}$ on a one time basis (payable in installments of $\$A_M$ annually over the next five years) to avoid these changes?

- ❑ NO ➔ If no, please continue to part c of this question
- ❑ YES➔ If yes, please move on to Plan B

c. Would you be willing to pay $\boldsymbol{\$B_L}$ on a one time basis (payable in installments of $\$A_L$ annually over the next five years) to avoid these changes?

- ❑ NO
- ❑ YES

A.3 Double-Bounded

1. If the current rates of drainage and sedimentation continue throughout the next decade, the water quality of Storm Lake will change to:

Fish catch per hour (May/June):	**1 fish every 4 hours**
Average lake depth:	**7 feet**
Muck on lake bottom:	**20 inches**
Water clarity :	**Objects distinguishable ½ foot under water**

a. Would you be willing to pay $\boldsymbol{\$B_M}$ on a one time basis (payable in installments of $\$A_M$ annually over the next five years) to avoid these changes and maintain Storm Lakes' current condition?

- ❑ NO ➔ If no, please continue to part b of this question
- ❑ YES➔ If yes, please move on to part c of this question

b. Would you be willing to pay $\boldsymbol{\$B_L}$ on a one time basis (payable in installments of $\$A_L$ annually over the next five years) to avoid these changes?

- ❑ NO
- ❑ YES

➔ Please move on to Plan B

c. Would you be willing to pay $\boldsymbol{\$B_H}$ on a one time basis (payable in installments of $\$A_H$ annually over the next five years) to avoid these changes?

- ❑ NO
- ❑ YES

➔ Please move on to Plan B

APPENDIX B: LIKELIHOOD FUNCTIONS

The log-likelihood functions used in estimating the three models presented in Table 3 follow directly from equation (15) and the assumption that $\ln\varepsilon \sim N(0,\sigma)$. We consider each model in turn.

B.1 First Bid Only

The first bid only model corresponds to the traditional single-bounded dichotomous choice model. Given the model specifications in subsection 4.2, the i^{th} subject's response to the first bid offered reveals the dichotomous choice variable:

$$\tilde{d}_{1i}(b_{1i}) = \begin{cases} 1 & \ln\varepsilon_i \geq \ln b_{1i} - \alpha \\ 0 & \ln\varepsilon_i < \ln b_{1i} - \alpha \end{cases} \tag{B.1}$$

Let $\Phi(x)$ denote the c.d.f. for a standard normal distribution. The log-likelihood function for our model, using only the first bid responses, is then given by:

$$\ell = \sum_{i=1}^{n}\left\{\tilde{d}_{1i}\left(1-\Phi[(\ln b_{1i}-\alpha)/\sigma]\right)+(1-\tilde{d}_{1i})\Phi[(\ln b_{1i}-\alpha)/\sigma]\right\} \tag{B.2}$$

B.1 With Follow-up

Hanemann, Loomis and Kanninen (1991) provide a specification for the log-likelihood function in the case of the double-bounded follow-up format. The log-likelihood functions for the follow-up models is this paper are more complex, not only because of we model the anchoring effect, but also because of the variation in follow-up formats used in the Storm Lake surveys. Given the model specifications in subsection 4.2, the i^{th} subject's response to the k^{th} bid offered reveals the dichotomous choice variable:

$$\tilde{d}_{ki}(b_{ki}) = \begin{cases} 1 & \ln\varepsilon_i \geq \dfrac{\ln b_{ki} - \gamma\ln b_{1i}}{(1-\gamma)} - \alpha \\ 0 & \ln\varepsilon_i < \dfrac{\ln b_{ki} - \gamma\ln b_{1i}}{(1-\gamma)} - \alpha \end{cases} \tag{B.3}$$

Examining equation (B.3), it is clear that the sequence of $\tilde{d}_{ki}$'s places an individual's realization of $\ln(\varepsilon_i)$ into one of four disjoint regions of the real number line: $S_1 \equiv [-\infty, \tilde{\beta}_L - \alpha)$, $S_2 \equiv [\tilde{\beta}_L - \alpha, \tilde{\beta}_M - \alpha)$, $S_3 \equiv [\tilde{\beta}_M - \alpha, \tilde{\beta}_H - \alpha)$, or $S_4 \equiv [\tilde{\beta}_H - \alpha, +\infty)$, where

$$\tilde{\beta}_k \equiv \frac{\ln B_k - \gamma \ln b_1}{(1-\gamma)} \qquad k = L, M, H \qquad \text{(B.4)}$$

In particular, for the one-way up format:

$$\ln \varepsilon_i \in \begin{cases} S_1(-\infty, \tilde{\beta}_L - \alpha) & \tilde{d}_1(\tilde{\beta}_L) = 0 \\ S_2[\tilde{\beta}_L - \alpha, \tilde{\beta}_M - \alpha) & \tilde{d}_1(\tilde{\beta}_L) = 1 \text{ and } \tilde{d}_2(\tilde{\beta}_M) = 0 \\ S_3[\tilde{\beta}_M - \alpha, \tilde{\beta}_H - \alpha) & \tilde{d}_1(\tilde{\beta}_L) = \tilde{d}_2(\tilde{\beta}_M) = 1 \text{ and } \tilde{d}_3(\tilde{\beta}_H) = 0 \\ S_4[\tilde{\beta}_H - \alpha, +\infty) & \tilde{d}_1(\tilde{\beta}_L) = \tilde{d}_2(\tilde{\beta}_M) = \tilde{d}_3(\tilde{\beta}_H) = 1 \end{cases} \qquad \text{(B.5)}$$

Similarly, for the one-way up format

$$\ln \varepsilon_i \in \begin{cases} S_1(-\infty, \tilde{\beta}_L - \alpha) & \tilde{d}_1(\tilde{\beta}_H) = \tilde{d}_2(\tilde{\beta}_M) = \tilde{d}_3(\tilde{\beta}_L) = 0 \\ S_2[\tilde{\beta}_L - \alpha, \tilde{\beta}_M - \alpha) & \tilde{d}_1(\tilde{\beta}_H) = \tilde{d}_2(\tilde{\beta}_M) = 0 \text{ and } \tilde{d}_3(\tilde{\beta}_L) = 1 \\ S_3[\tilde{\beta}_M - \alpha, \tilde{\beta}_H - \alpha) & \tilde{d}_1(\tilde{\beta}_H) = 0 \text{ and } \tilde{d}_2(\tilde{\beta}_M) = 1 \\ S_4[\tilde{\beta}_H - \alpha, +\infty) & \tilde{d}_1(\tilde{\beta}_H) = 1 \end{cases} \qquad \text{(B.6)}$$

Finally, for the double-bounded format, we have:

$$\ln \varepsilon_i \in \begin{cases} S_1(-\infty, \tilde{\beta}_L - \alpha) & \tilde{d}_1(\tilde{\beta}_M) = 0 \text{ and } \tilde{d}_2(\tilde{\beta}_L) = 0 \\ S_2[\tilde{\beta}_L - \alpha, \tilde{\beta}_M - \alpha) & \tilde{d}_1(\tilde{\beta}_M) = 0 \text{ and } \tilde{d}_2(\tilde{\beta}_L) = 1 \\ S_3[\tilde{\beta}_M - \alpha, \tilde{\beta}_H - \alpha) & \tilde{d}_1(\tilde{\beta}_M) = 1 \text{ and } \tilde{d}_2(\tilde{\beta}_H) = 0 \\ S_4[\tilde{\beta}_H - \alpha, +\infty) & \tilde{d}_1(\tilde{\beta}_M) = 1 \text{ and } \tilde{d}_2(\tilde{\beta}_H) = 1 \end{cases} \qquad \text{(B.7)}$$

Let

$$I_{ki}(x, y) \equiv \begin{cases} 1 & \ln \varepsilon_i \in S_k(x, y) \\ 0 & \text{otherwise} \end{cases} \qquad \text{(B.8)}$$

The log-likelihood for the our basic model, when the follow-up questions are used and anchoring is ignored, is then given by:

$$\begin{aligned}\ell = \sum_{i=1}^{n} \{ & I_{1i} \ln\left(\Phi[(\ln B_{Li} - \alpha)/\sigma]\right) \\ & + I_{2i} \ln\left(\Phi[(\ln B_{Mi} - \alpha)/\sigma] - \Phi[(\ln B_{Li} - \alpha)/\sigma]\right) \\ & + I_{3i} \ln\left(\Phi[(\ln B_{Hi} - \alpha)/\sigma] - \Phi[(\ln B_{Mi} - \alpha)/\sigma]\right) \\ & + I_{4i} \ln\left(1 - \Phi[(\ln b_{Hi} - \alpha)/\sigma]\right)\}\end{aligned} \tag{B.10}$$

When anchoring is controlled for, the likelihood function changes to:

$$\begin{aligned}\ell = \sum_{i=1}^{n} \{ & I_{1i} \ln\left(\Phi[(\{[\ln B_{Li} - \gamma \ln b_{1i}]/(1-\gamma)\} - \alpha)/\sigma]\right) \\ & + I_{2i} \ln\left(\Phi[(\{[\ln B_{Mi} - \gamma \ln b_{1i}]/(1-\gamma)\} - \alpha)/\sigma] - \Phi[(\{[\ln B_{Li} - \gamma \ln b_{1i}]/(1-\gamma)\} - \alpha)/\sigma]\right) \\ & + I_{3i} \ln\left(\Phi[(\{[\ln B_{Hi} - \gamma \ln b_{1i}]/(1-\gamma)\} - \alpha)/\sigma] - \Phi[(\{[\ln B_{Mi} - \gamma \ln b_{1i}]/(1-\gamma)\} - \alpha)/\sigma]\right) \\ & + I_{4i} \ln\left(1 - \Phi[(\{[\ln B_{Hi} - \gamma \ln b_{1i}]/(1-\gamma)\} - \alpha)/\sigma]\right)\}\end{aligned} \tag{B.11}$$

REFERENCES

Alberini, A., B. Kanninen, and R. Carson, "Estimating Willingness to Pay with Discrete Choice Contingent Valuation Data Using Error Component Models," Presented at the AERE session of the ASSA Meetings, Boston, 1994.

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman, "Report of the NOAA Panel on Contingent Valuation." Federal Register 58(1993): 4601-4614.

Bishop, R., and T. Heberlein, "Measuring Values of Extra-Market Goods: Are Indirect Measures Biased?" *Amer. J. Agr. Econ.* 61(1979): 926-930.

Bishop, R., and T. Heberlein, "Simulated Markets, Hypothetical Markets, and Travel Cost Analysis: Alternative Methods of Estimating Outdoor Recreation Demand." Staff Paper Series no. 187, Department of Agricultural Economics, University of Wisconsin, 1980.

Boyle, K., R. Bishop, and M. Welsh, "Starting Point Bias in Contingent Valuation Surveys." *Land Economics* 61(1985): 188-194.

Cameron, T., and M. James, "Efficient Estimation Methods for Use with 'Closed-Ended' Contingent Valuation Survey Data." *Rev. Econ. and Statist.* 69(1987): 269-76.

Cameron, T., and J. Quiggin, "Estimation Using Contingent Valuation Data from a 'Dichotomous Choice with Follow-up' Questionnaire." *J. Environ. Econ. and Manage.* (1993), forthcoming.

Carson, R., "Three Essays on Contingent Valuation." Ph.D. thesis, University of California, Berkeley, 1985.

Carson, R., W. Hanemann, and R. Mitchell, "Determining the Demand for Public Goods by Simulating Referendums at Different Tax Prices." Dep. Econ. Work. Pap., University of California, San Diego, 1986.

Combs, J., R. Kirkpatrick, J. Shogren, and J. Herriges, "Matching Grants and Public Goods: A Closed-Ended Contingent Valuation Experiment." *Public Finance Quarterly* 21(1993): 178-195.

Cooper, J., "Optimal Bid Selection for Dichotomous Choice Contingent Valuation Surveys." *J. Environ. Econ. Manage.* 24(1993): 25-40.

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

Hanemann, W., "Some Issues in Continuous- and Discrete-Response Contingent Valuation Studies." *Northeast. J. Agr. Econ.* (1985):5-13.

Hanemann, W., J. Loomis, and B. Kanninen, "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation." *Amer. J. Agr. Econ.* 73 (1991):1255-1263.

Jones, J., and R. Bachmann, "Trophic Status of Iowa Lakes in Relation to Origin and Glacial Geology." *Hydrobiologia* 57(1978): 267-273.

Kanninen, B., "Design of Sequential Experiments for Contingent Valuation Studies." *J. Environ. Econ.and Manage.* 25(1993a): S1-S11.

Kanninen, B., "Optimal Experimental Design for Double-Bounded Dichotomous Choice Contingent Valuation." *Land Economics* 69(1993b): 138-146.

McFadden, D., and G. Leonard, “Issues in the Contingent Valuation fo Environmental Goods: Methodologies for Data Collection and Analysis.” in *Contingent Valuation: A Critical Assessment*, J. A. Hausman, ed., North Holland, Amsterdam.

Menzel, B., and R. Schultz, “A Progress Report and Renewal Request Prepared for the Leopold Center for Sustainable Agriculture.” mimeograph, March 1993.

Mitchell, R., and R. Carson, *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington DC: Resources for the Future, 1989.

Table 1
Experimental Design

Follow-up	Format	Target Population	
		Local Residents	Recreationists
No	Single-Bounded	0	300
Yes	One-Way Street Up	100	100
Yes	Double-Bounded	100	100
Yes	One-Way Street Down	100	100

Table 2
Bid Values

a. Local Residents

Sample	B_L	B_M	B_H
150	100	200	300
150	200	300	400
Mean Value	150	250	350

b. Recreationists[24]

Sample	B_L	B_M	B_H
75	50	125	250
75	75	150	275
75	100	175	325
75	125	225	475
Mean Value	87.5	168.75	331.25

[24] The sample sizes listed here refer to the 300 surveys using follow-up. The surveys without follow-up followed exactly their follow-up counter-parts, but simply excluded the follow-up questions.

Table 3
Parameter Estimates

	Local Residents			Recreationists		
	First Bid Only	With Follow-up		First Bid Only	With Follow-up	
		γ=0	γ Unconstrained		γ=0	γ Unconstrained
Median WTP: exp(α)	148.85 (20.1)	139.75 (13.30)	140.65 (17.80)	41.90 (15.70)	64.70 (7.45)	43.10 (13.80)
WTP Dispersion: σ	1.14 (.26)	1.12 (.12)	1.10 (.23)	2.35 (.61)	1.48 (.15)	2.10 (.46)
Anchoring Effect γ	n.a.	0	.00 (.20)	n.a.	0	.36 (.14)

Figure 1
Alternative Follow-up Formats

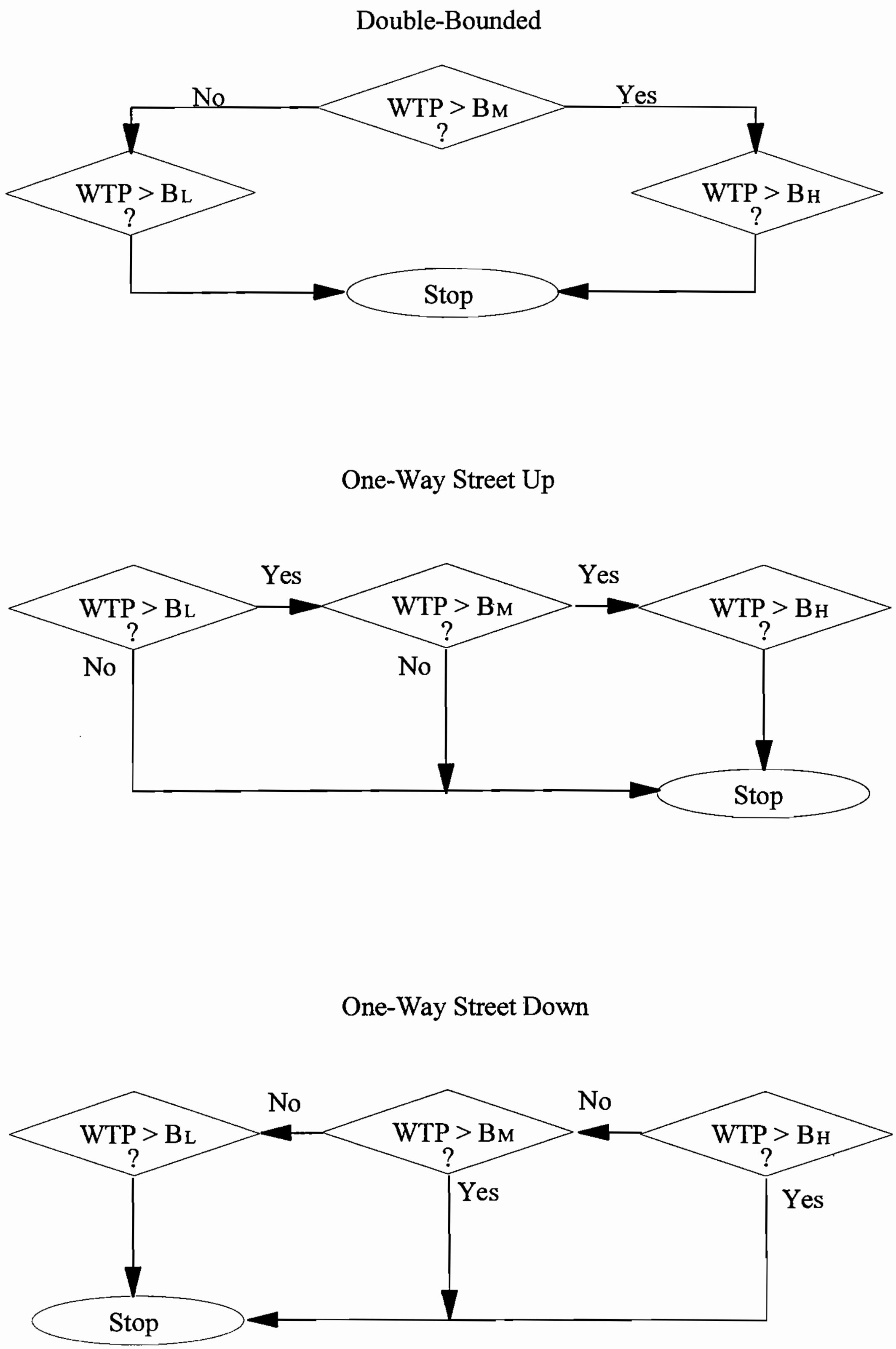


Figure 2
Response Rates

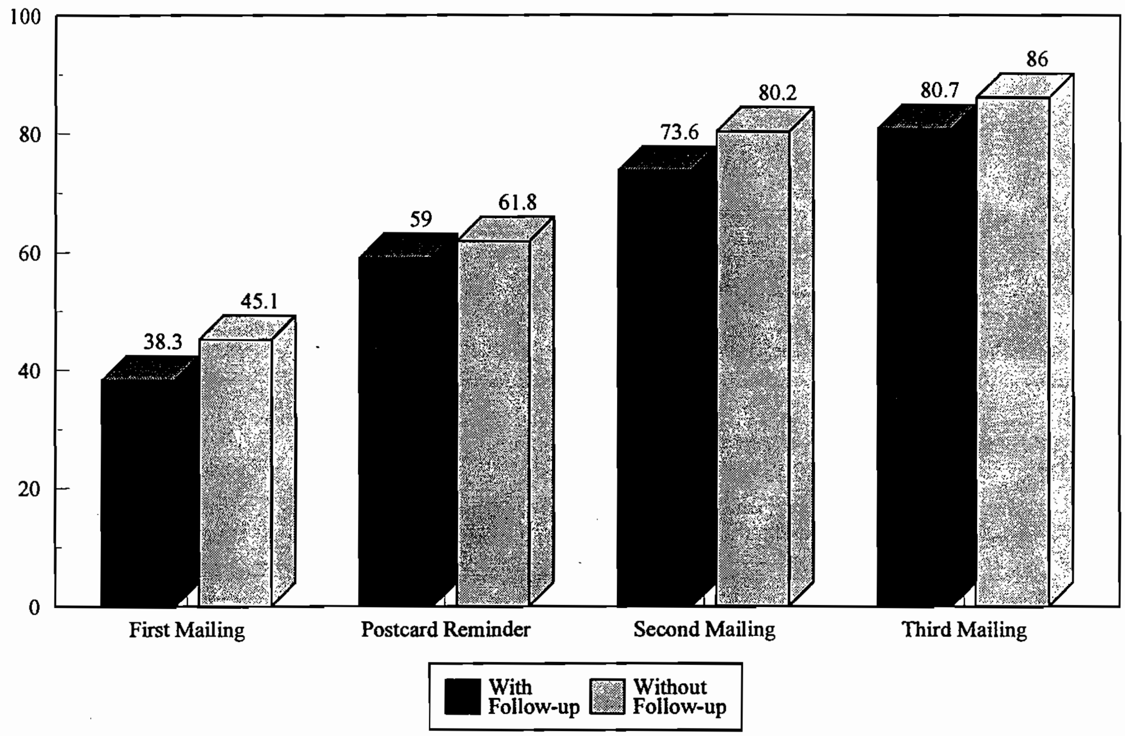
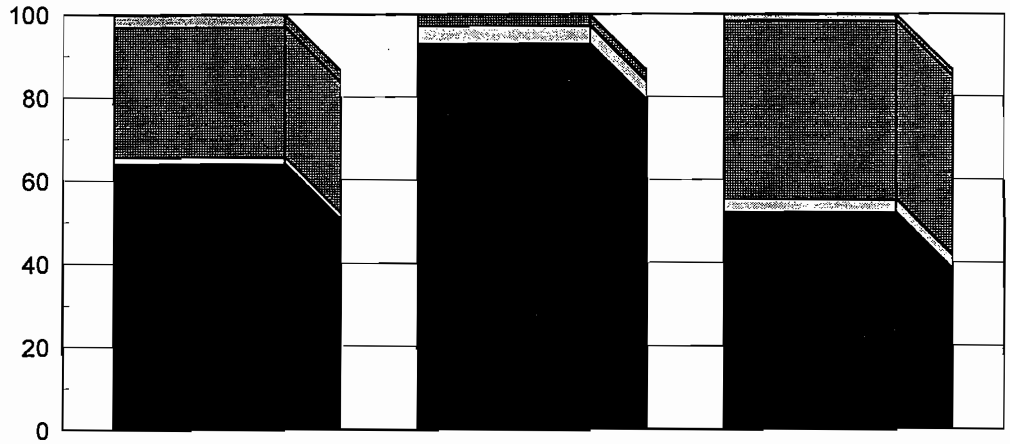


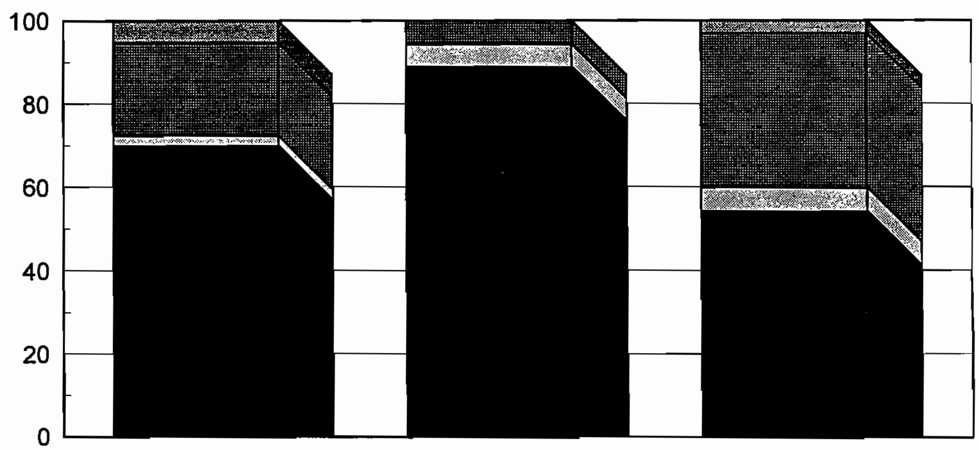
Figure 3
Recreationists Response Patterns

Total Sample



	One-Way Up	One-Way Down	Double-Bounded
Complete	64.2	93.0	52.3
Incomplete	1.5	4.2	3.1
Overkill	31.3	2.8	43.1
Inconsistent	3.0	0.0	1.5

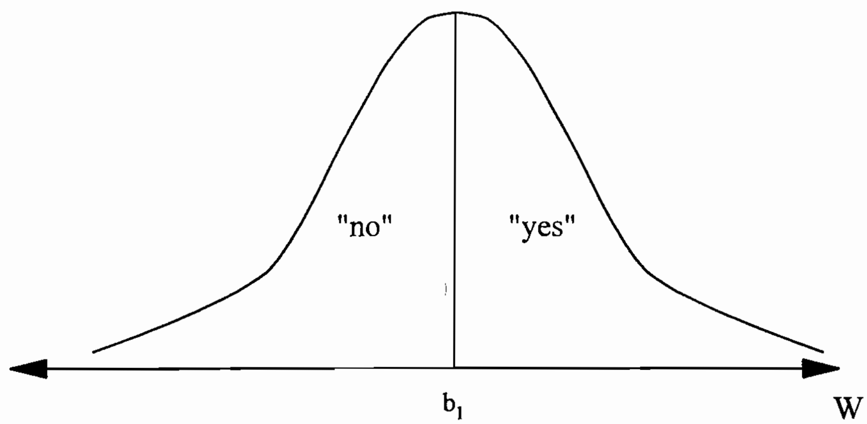
Excluding Protests



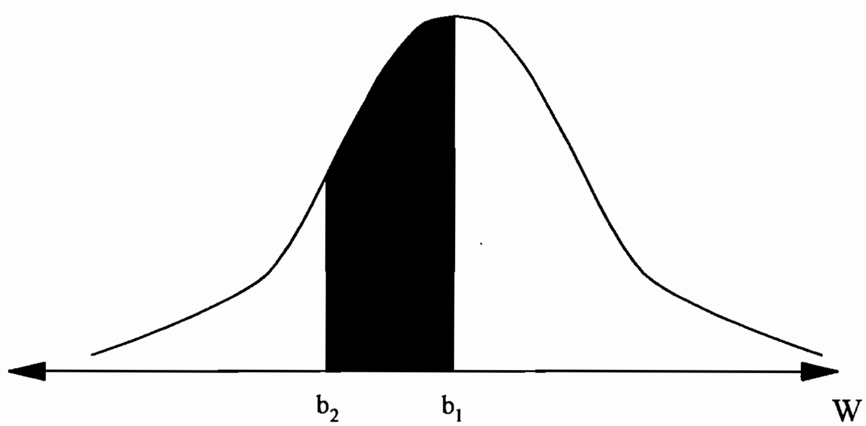
	One-Way Up	One-Way Down	Double-Bounded
Complete	70.0	89.0	54.3
Incomplete	2.5	5.5	5.7
Overkill	22.5	5.5	37.1
Inconsistent	5.0	0.0	2.9

Figure 4
Dichotomous Choice Bounds

a. Single Bounded



b. Double- Bounded - No Anchoring



b. Double- Bounded - No Anchoring

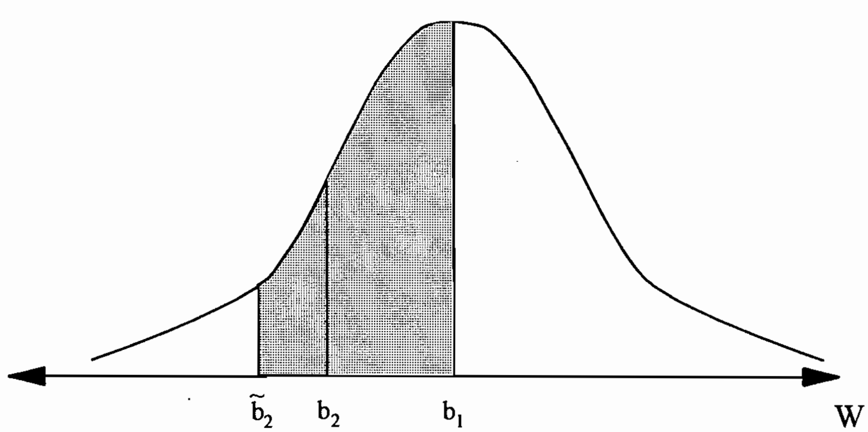
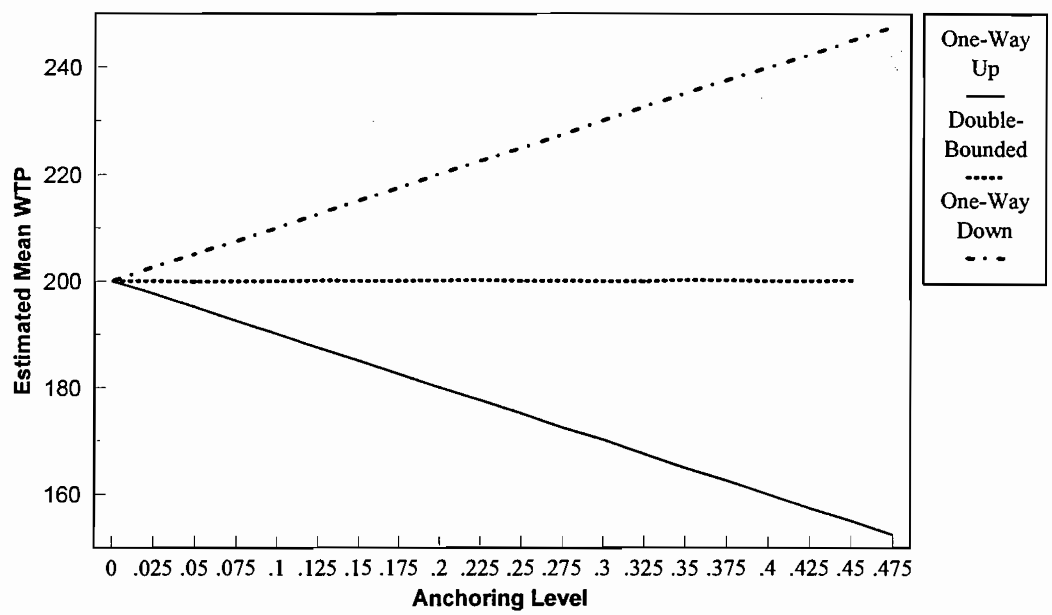
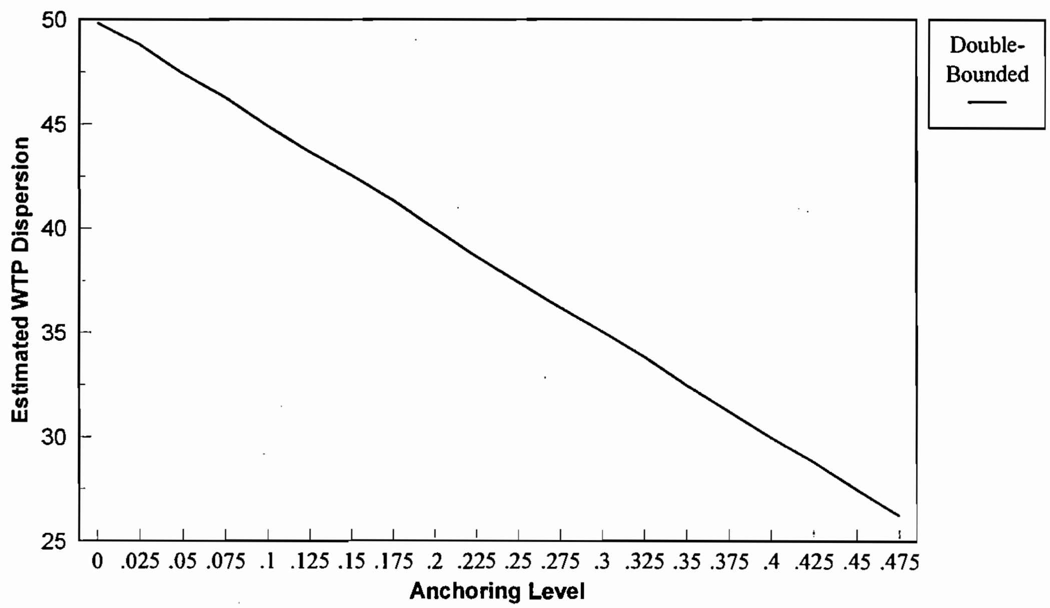


Figure 5
Simulation Results

Mean WTP



WTP Dispersion Estimate



Natural Resource Damage Assessment under the Oil Pollution Act:
Issues for Comment in NOAA's Proposed Regulations

Carol Adaire Jones
NOAA Damage Assessment Center
1305 EastWest Highway Station 10218
Silver Spring MD 20910-3281
Internet: cjones@nos.noaa.gov

On January 7, 1994 NOAA proposed regulations for performing natural resource damage assessments under the Oil Pollution Act of 1990. [59 FR 1062-1191, (Part II]. The period for receiving public comments closes on July 7, 1994. Below is a discussion of issues pertaining to the calculation of interim lost values pending recovery of injured resources/services, also known as 'compensable values', on which NOAA is seeking comment.

COMPENSABLE VALUES ISSUES (OTHER THAN CONTINGENT VALUATION)

Much attention has been directed on proposed methods for valuing damages to natural resources resulting from discharges of oil, particularly the controversial contingent valuation method. However, given the expense and time associated with designing and implementing site-specific studies -- employing contingent valuation or other methods -- natural resource trustees have employed alternative methods for the majority of oil spill damage assessments.

This section discusses and seeks comment on two methods included in the natural resource damage assessment regulations promulgated pursuant to the Oil Pollution Act of 1990: the benefits transfer method and the habitat or species replacement cost method. These alternatives allow the trustee to estimate damages for interim lost services at a lower cost than with site-specific studies. Also, a third alternative may be considered that would involve government expenditures on a resource as a lower bound for the "value" of that resource. A fourth related issue that we seek comment on is the appropriate discount rate to be used in calculating the present value of natural resource damage claims.

Benefits Transfer: The benefits transfer method uses existing estimates of use values or of valuation functions that were developed in one context to address a similar resource valuation question in a different context. Transfer of these resource or use values may allow a less time-consuming and less expensive damage estimation than site-specific valuation analysis.

Specific questions on which NOAA is seeking comment:

In the preamble discussion, NOAA has proposed the following criteria for trustees to consider in determining whether benefits transfer is appropriate, identifying several questions a trustee(s) should consider for each criterion:

- the comparability of the resources/services at the original study and at the particular injury assessment site (the transfer site);
- the comparability of the change in the quality or quantity of the resources/services at the original study site and the particular assessment site; and
- the quality of the original study.

NOAA is seeking comments on whether these are the appropriate factors to consider, and whether there are other factors that should be considered.

Additional questions on which NOAA is seeking comment pertain to the use of contingent valuation (CV) studies in a benefits transfer context.

- What conditions should be placed on the transfer of CV studies? Should the proposed requirements for site-specific CV studies measuring passive use value serve a role in determining whether the quality of a CV study is adequate for use in benefits transfer? If so, what role should they serve?

- Should the conditions be different for valuing direct use losses and for valuing passive use losses?
- How would the extent of the market be determined for benefits transfer using CVs?

Habitat or Species Replacement Cost: The habitat or species replacement cost method may be used to estimate damages for lost services from the injured habitats and/or biological resources, when human services provided by the habitat or species are difficult to quantify. This method involves estimating damages in terms of the cost of obtaining from an alternative source the equivalent of the interim lost value of resources and/or services. The method will be utility-theoretic, if the flow of services provided by the replacement projects provide an gain in (lifetime) utility comparable to that lost by the injuries.

In order to ensure that the scale of the compensatory restoration or replacement project(s) on which the cost calculation is based does not over- or under-compensate the public for injuries incurred, the proposed OPA rule suggests that the trustee must establish an equivalency between the present discounted value (PDV) of the quantity of lost services and the PDV of the quantity of services provided by the replacement project(s) over time.

Specific questions on which NOAA is seeking comment:

NOAA is specifically seeking comment on whether the proposed OPA rule has suggested reasonable requirements for the use of this method. Are there other conditions that must, or should, be considered?

Government Expenditures: Participants in public meetings have proposed the use of government expenditures (per unit of services) as a proxy for the value of those services lost as a result of an incident.

Specific questions on which NOAA is seeking comment:

- Should the OPA rule suggest that government expenditures may be used as a proxy for the value of a resource?
- Under what circumstances would use of this method be appropriate?

For example, two conditions could be that(1)the changes in the quality/quantity of the injured resources should be related to the (change in) level of government expenditures, and (2) the government programs for which expenditures have been accounted are the major cause of the changes in the quality/quantity of the affected resource.

Discount Rate: The proposed OPA rule specifies that the trustee is to use the U.S. Treasury rate to discount all three categories of damages: restoration costs, interim lost value, and damage assessment costs. Following the guidance in OMB Circular A-94 (for cost-effectiveness analysis and for federal leasing), nominal interest rates are to be used with damages in nominal terms (for example, past damage assessment costs) and real rates (with an adjustment based on the Administration's prediction for future inflation published in the President's budget) are to be used for damages in real terms.

Specific questions on which NOAA is seeking comment:

For restoration costs, what if the return on accounts available to the trustee(s) for placement of recovered funds is lower than the U.S. Treasury rate, so that the present discounted value of future restoration costs will not support the full restoration project?

CONTINGENT VALUATION

NOAA is proposing that contingent valuation (CV) can produce estimates of interim lost value, including lost passive use value, that are reliable enough for use in a judicial or administrative determination of natural resource damages, including lost passive use value. To achieve the reliability necessary for this purpose, the studies must adhere to a set of requirements outlined in the proposed regulations. This finding follows the guidance provided by the report of the NOAA Blue Ribbon Panel on Contingent Valuation (chaired by Professors Kenneth Arrow and Robert Solow) and other comments submitted to NOAA.

The four categories of requirements included in the proposed OPA rule are:

- Survey Design: Requirements specific to CV surveys
- Survey Administration: Generic requirements to produce reliable surveys
- Nature of the Results: Internal validity checks on the survey results
- Reporting: Complete reporting of survey instrument, data, and analysis, plus documentation of rationale for design choices

NOAA worked closely with The Department of the Interior in drafting and refining the contingent valuation (CV) language for the rule and in outlining the scope of a possible CV guidance document. The Department of the Interior and NOAA agreed to propose similar language for both the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) and Oil Pollution Act (OPA) rules, thereby providing consistent approaches to CV in the two regulatory schemes.

Regulatory language on CV on which NOAA is seeking comments:

Calibration of elicited values (comparability with "real" transactions): The proposed OPA rule provides that the respondents' stated values be divided by two, unless trustee can justify an alternative calibration factor for the specific case. This calibration procedure is to "correct" for the combined effects of two countervailing potential biases: the mandated elicitation of willingness to pay (WTP) measures may understate the correct measure of damages [willingness to accept], whereas the elicitation of hypothetical WTP in contingent valuation studies may overstate "true" WTP.

Is calibration appropriate? If so, is the proposed default calibration factor appropriate? On what basis could "calibration" factors be developed for individual cases?

Performance Tests: Internal Validity Checks: The proposed OPA rule states that two independent "scope" tests are to be conducted. The tests must show significant changes in respondents' WTP in response to variations in the scope of injuries (from the injury scenario to be proved in the case), unless the trustee(s) can show that conducting the test with two scenarios in addition to the base case is infeasible due to considerations of cost or lack of plausibility of scenarios. The tests are to be conducted with a separate sample for each scenario, (a "split sample" test). Proposed regulations outline procedures to limit the differences between the base case and alternative scenarios in order to strengthen the internal validity tests.

Scope of injury testing: How many sensitivity to "scope of injury" tests are appropriate to require? Are restrictions on differences between scenarios appropriate? Feasible?

Alternative testing methods: Is it appropriate to allow the trustee(s) to conduct the scope test with the base survey instrument, by constructing a valuation function to examine whether variations in belief about injuries predict variations in WTP, controlling for demographic and attitudinal factors? Are there additional/alternative internal validity test(s) that NOAA should consider?

Response rate: 70% is the minimum allowable response rate contained in the proposed OPA rule. In order to minimize non-response bias, should a minimum response rate be specified? If so, is 70% the appropriate level? If 70% is not the appropriate rate, what rationale is there for a different rate?

Characteristics of choice mechanism and payment vehicle: In the proposed OPA rule, the trustee(s) is directed to use a choice mechanism that is credible and incentive-compatible, i.e., one that does not provide respondents with incentives to understate or overstate their true value. The reasons are outlined in the preamble for recommending the use of a referendum as the choice mechanism in a survey. Are these requirements appropriate?

Additional Issues Not in the Proposed Rule, but on which NOAA is Seeking Comments

Role of prior information about resources/injuries: Should respondents with no knowledge of the resources and/or injuries prior to survey be assigned a zero value for damages due to the injuries? What is the appropriate use of data on respondents' prior information?

Screening test to limit use of CV to "high value" cases: Should there be thresholds for damages, below which CV could not be used in a damage assessment, e.g., an expected $5/household times the number of households expected to hold passive values; and/or twice the cost of a contingent valuation survey following these regulations. What threshold, if any, is appropriate? How would the threshold be implemented (without performing a CV study)?

Mode of administration: In the proposed rule, the trustee(s) has the option of choosing the mode of administration of a survey, but the choice must be justified. Is one mode, e.g., in-person, telephone, or mail, preferable to another? What is the rationale?

Extent of the market for passive use values: Should the rule or preamble provide guidance on criteria for determining the extent of the market? If yes, what criteria would be appropriate?

Requirements for CV studies to measure direct use values only: What requirements, if any, should be imposed on CV studies for valuing direct use only?

Requirements for contingent behavior studies: What requirements, if any, should be imposed on contingent behavior studies?

Send all comments, or requests for copies of the rule and supporting documentation (preferably on disk), to:

NOAA Damage Assessment and Regulations Team
202-606-8000; 202-606-4900 FAX
c/o NOAA Damage Assessment Center
1305 East West Highway Station 10218
Silver Spring MD 20910-3281

EFFICIENCY v. BIAS OF WILLINGNESS-TO-PAY ESTIMATES:

BIVARIATE AND INTERVAL-DATA MODELS.

by

Anna Alberini[1]
Resources for the Future
1616 P Street, NW
Washington, DC 20036
(e-mail: alberini@rff.bitnet)

Abstract

Dichotomous choice contingent valuation (CV) surveys with a follow-up attempt to obtain more precise information about the respondents' willingness to pay for environmental quality while at the same time preserving the desirable properties of the referendum CV method. This paper offers a contribution along this line of research by comparing the interval-data models traditionally associated with these surveys and the recently proposed *bivariate* binary response models. The latter allow for the responses to the two payment questions to be driven by two, potentially different, and potentially correlated, unobserved willingness to pay (WTP) values. A simulation study shows that the interval-data WTP estimates can be surprisingly robust to departures from the true, bivariate model, and that in many situations with highly correlated WTP values, which are likely to be of interest to CV researchers, the interval-data model is superior to the bivariate model for WTP in terms of the mean square error of the estimates. This is confirmed by one of the two empirical examples reported in the paper, in which the interval-data estimates might be preferred over the bivariate model estimates because of their small biases and large gains in efficiency. Another example demonstrates that the two competing models have a different degree of robustness in the presence of observations in the tails of the WTP distribution, a poor choice of the fitted WTP distribution, and responses contaminated by strategic considerations or "yea-saying".

---

[1]Anna Alberini is a Fellow at Resources for the Future, Quality of the Environment Division. The Author would like to thank Richard Carson, Barbara Kanninen and Ray Kopp for their comments on earlier versions of the paper. The Author gratefully acknowledges the support of the Fondazione E.N.I. Mattei, Milan (Italy). The usual disclaimer applies.

EFFICIENCY v. BIAS OF WILLINGNESS-TO-PAY ESTIMATES: BIVARIATE AND INTERVAL-DATA MODELS.

1. INTRODUCTION.

The discrete choice approach to eliciting willingness to pay (WTP) in contingent valuation surveys has become a commonly used method to obtain information about the public's valuation of a good, environmental plan or natural resource. The NOAA Panel on Contingent Valuation [2] has recently sanctioned the use of the dichotomous choice approach to contingent valution (CV), and thus consolidated its popularity. Researchers are by now familiar with fitting logit, probit or other binary response models to discrete choice CV survey data to obtain estimates of the parameters of the latent WTP variable, and statistics that are meaningful for benefit-cost analysis purposes, such as mean or median WTP.

In order to refine the information about WTP while at the same time preserving the desirable characteristics of discrete choice surveys, CV practitioners have implemented discrete choice surveys with a *follow-up* payment question (Carson, Hanemann and Mitchell [4]). The corresponding double-bound or interval-data models of WTP have been shown to produce more efficient estimates than those obtained fitting single-bound models of WTP that use only the responses to the first payment question (Hanemann, Loomis and Kanninen [7]).

Much controversy has, however, been stirred by the extreme sensitivity of interval-data WTP estimates to the choice of the distribution for WTP (see Carson et al., [5], for an example), and by reports of CV surveys in which WTP was found to change systematically between the first and the second payment questions (Hanemann, Loomis and Kanninen [7]; McFadden and Leonard [9]). Cameron and Quiggin [3] recently suggested that respondents might form *two* WTP values, one for each of the payment questions, and that these two WTP values are likely to be correlated but need not be identical. The *bivariate* binary response model derived under this assumption was applied to the data from a CV survey conducted in Australia and resulted in estimates of WTP from a bivariate model that were dramatically different from the double-bound estimates, even though the double-bound model was expected to be a rather close approximation of the bivariate model. Since the WTP estimates from CV surveys are used to form policy recommendations, and to recover the damages to natural resources caused by accidents involving hazardous substances and oil spills under CERCLA (1980), these findings are disturbing. They seem to imply that small changes in the econometric models of WTP or in the WTP elicitation technique, such as addding a follow-up question, can refute the feasibility of a policy, or the amount of the settlement for the damage to natural resources sought by government agencies in CERCLA litigations.

This paper focuses on the comparison between the WTP estimates from double-bound and bivariate models based on a normal WTP. In Section 2 it is shown that the standard double-bound model is a special case of a bivariate probit model in which appropriate restrictions are imposed on the parameters of the WTP to ensure that the first and second WTP values are identical. Among the other restrictions, a double-bound model assumes that the correlation between the first and the second WTP amounts is perfect. Section 3 presents a simulation exercise that explores the effect on the estimates of fitting an interval-data model although a bivariate probit model is the correct statistical framework, and viceversa. As expected, the interval-data estimates are always more efficient, but become biased when the true correlation coefficient departs from one. For sufficiently high correlation, however, the biases are very small relative to the gains in efficiency ensured by the interval-data model, and the latter is superior in terms of the mean square errors of the estimates. CV practitioners are further reminded that, while it might be tempting to estimate a bivariate probit model and check the interval-data specification using a classical test of the null hypothesis that the correlation coefficient is one, it is not correct to proceed in this fashion. The classical tests have been developed under the assumption that the true value of the parameter, or the value hypothesized by the null, lies in an *interior* of the parameter space (*i.e.*, the correlation coefficient ρ lies strictly within the (-1,1) range), and their properties when the true parameter ρ is on the frontier of the parameter space ($\rho=\pm1$) are not well known.

Two empirical examples are reported in Section 4 that demonstrate how a model selection strategy might be based on trading off efficiency and bias of the estimates, and how large discrepancies between the interval-data and the bivariate probit estimates may due to a non-normal, thick-tailed WTP, or possibly to the failure of the maintained mapping between latent WTP's and observed responses. Section 5 provides some concluding remarks.

2. INTERVAL-DATA AND BIVARIATE PROBIT MODELS.

Interval-data models assume that one respondent's responses to the initial and the follow-up dichotomous choice payment questions are driven by a single WTP amount. The survey protocol itself is based on this assumption, since the bid offered with the second question depends on the information gained from the response to the initial payment question. The contributions to the likelihood of an interval-data model from each pair of observations are given in Table I for the situation in which $WTP_i = x_i\beta + \varepsilon_i$, where ε_i is a normal i.i.d. error term with mean zero and variance σ^2, x_i is a vector of individual characteristics and β is a vector of unknown parameters of compatible dimension.

Table I. Contributions to the likelihood for bivariate probit and interval-data probit model.[a]

type of responses	interval-data probit (1)	bivariate probit (2)
no,no	$\Phi(z_{i2})$	$\Phi(z_{i1}, z_{i2}, \rho)$
no,yes	$\Phi(z_{i1}) - \Phi(z_{i2})$	$\Phi(z_{i1}) - \Phi(z_{i1}, z_{i2}, \rho)$
yes,no	$\Phi(z_{i2}) - \Phi(z_{i1})$	$\Phi(z_{i2}) - \Phi(z_{i1}, z_{i2}, \rho)$
yes,yes	$1 - \Phi(z_{i2})$	$1 - \Phi(z_{i1}, z_{i2}, \rho) - \Phi(z_{i1}) - \Phi(z_{i1})$

[a] $z_{i1} = (c_{i1} - x_i\beta_1)/\sigma_1$, $z_{i2} = (c_{i2} - x_i\beta_2)/\sigma_2$, c_{i1} is the first bid assigned to the i-th respondent, c_{i2} is the follow-up bid, $\Phi(\bullet)$ is the standard normal cdf and $\Phi(\bullet,\bullet,\rho)$ is the standard bivariate normal cdf with correlation coefficient ρ. $z_{2i} < z_{1i}$ $(z_{2i} > z_{1i})$ if the answer to the initial payment question is no (yes).

Cameron and Quiggin [3] suggested that the respondents may well refer to *two* WTP values in the survey, one for each of the two payment questions. The underlying WTP values is described by the bivariate system:

$$(1) \qquad \begin{cases} WTP_{i1} = x_i\beta_1 + \varepsilon_{i1} \\ WTP_{i2} = x_i\beta_2 + \varepsilon_{i2} \end{cases} \quad .$$

The error terms ε_1 and ε_2 are assumed to be jointly normally distributed with means zero, variances σ_1^2 and σ_2^2, and correlation ρ. The bivariate model (1) can be interpreted to mean that, if the respondents' exact numerical WTP values could be observed, the first could be used to construct a linear and unbiased predictor for the second using the properties of conditional normal distributions: $E(WTP_{i2} | WTP_{i1}) = x_i\beta_2 + \frac{\rho\sigma_2}{\sigma_1}[WTP_{i1} - x_i\beta_1]$. Such prediction would, of course, be subject to a random prediction error: the higher the correlation between the two variables, the better the prediction (the smaller the variance of the prediction error). If both responses are driven by a single WTP amount, $\beta_1 \equiv \beta_2$, $\sigma_1 \equiv \sigma_2$, and ρ=1, the prediction error is equal to zero and $WTP_{i1} \equiv WTP_{i2}$. The contributions to the likelihood from each pair of responses are given in Table I. The familiar interval-data (or double-bound) model is a special case of the more general bivariate model with $\beta_1 \equiv \beta_2$, $\sigma_1 \equiv \sigma_2$, and $\rho \equiv 1$, which forces WTP_{i1} to be identical to WTP_{i2}, since $\Phi(z_{i1}, z_{i2}, \rho)$ tends to $\Phi(\min(z_{i1}, z_{i2}))$ for $\rho \to 1$.[2]

Specific hypotheses about the β's, σ's and the free correlation parameter ρ of the bivariate probit model can be tested by means of a classical test of hypotheses, such as a likelihood ratio or a Wald-type test.[3] Either test could be used, for instance, to check that $\beta_1 \equiv \beta_2$ and $\sigma_1 \equiv \sigma_2$, the choice of the test being dictated by convenience and computational ease, since they are asymptotically equivalent under the null

---

[2]Using the data from a CV survey about the use of natural resources of Kakadu Conservation Zone and Kakadu National Park in Australia (Carson, Wilks and Imber [6]), Cameron and Quiggin [3] fitted an interval-data model based on a normal WTP, obtaining an estimated mean/median WTP of about A$ 510. A bivariate probit model which relaxed the assumption that ρ=1 and estimated a free correlation parameter produced a dramatically different estimated mean/median WTP of A$ 150 and a correlation coefficient between individual WTP amounts of 0.95 (see Appendix).

[3]The classical tests of hypothesis rely on the asymptotic distribution of the parameter estimates under the null being normal and centered around the values specified by the null hypothesis. The likelihood ratio (LR) test is based on the difference between the constrained and the unconstrained log likelihood functions evaluated at their respective optima. The test statistic is LR=$-2*[\log L - \log L_0]$, where $\log L_0$ is the constrained log likelihood function. The Wald test requires only the unconstrained estimates and is generally written as $W = (\hat{\beta} - \beta_0)' V^{-1}(\hat{\beta})(\hat{\beta} - \beta_0)$, where $V(\hat{\beta})$ is the covariance matrix of $\hat{\beta}$ and β_0 is the value of β specified by the null hypothesis. The likelihood ratio, the Wald and another classical test, the score test (Amemiya [1]), are distributed as a chi square with g degrees of freedom under the null hypothesis, g being the number of independent restrictions specified by the null.

hypothesis. It should be emphasized, however, that it is *not* appropriate from the statistical point of view to apply a classical test of the hypothesis that $\rho=1$.

The asymptotic distribution of convergence of the classical tests under the null hypothesis are, in fact, derived under the assumption that the true parameters are in the *interior* of the parameter space (see Amemiya [1]). A correlation coefficient of one, which is on the frontier of the parameter space, invalidates the latter assumption: as a result, the parameter estimates may not be well-behaved, the distribution of convergence of the test statistics may be non-standard and the empirical sizes of the test statistics (the frequency of rejection of the null hypothesis that the estimate of a parameter is equal to the true value of the parameter) may be different from the nominal sizes of the tests (5% or 1%). As a consequence, the outcome of a test of the null that $\rho=1$ should *not* be used as a means of model selection between the bivariate and the double-bound models.

3. EFFICIENCY V. BIAS: EVIDENCE FROM A SIMULATION EXERCISE.

A simulation exercise was designed and run to verify the effect of fitting a bivariate probit model when the appropriate statistical framework is the interval-data probit model, and fitting an interval-data probit model when the true WTP is a bivariate normal. Numbers representing the underlying WTP amounts on a continuous scale were randomly drawn from a bivariate normal variable with correlation coefficient ρ and split into pairs of binary responses using predetermined sets of threshold values.[4] The correlation coefficient was varied over different sets of simulations (ρ=0.2, 0.6, 0.9 and 1.0).

[4]The binary response corresponding to the first payment question was assigned a value of one if the latent WTP amount was greater than the first bid, and of zero otherwise. A similar procedure was followed to create the binary response corresponding the second payment question using the second WTP amount and the follow-up bid value. The set of first bids was comprised of the values ±1.35, $\pm0.91\sigma+\mu$, $\pm0.60\sigma+\mu$, $\pm 0.25\sigma+\mu$ and $\pm0.11\sigma+\mu$. The set of the second bids of the values included $0.11\sigma+\mu$, $0.23\sigma+\mu$, $0.35\sigma+\mu$, $0.47\sigma+\mu$, $0.60\sigma+\mu$, $0.75\sigma+\mu$, $0.91\sigma+\mu$, $1.10\sigma+\mu$, $1.33\sigma+\mu$, $1.69\sigma+\mu$ (if the first response was positive) and $-1.69\sigma+\mu$, $-1.33\sigma+\mu$, $-1.10\sigma+\mu$, $-0.91\sigma+\mu$, $-0.75\sigma+\mu$, $-0.60\sigma+\mu$, $-0.47\sigma+\mu$, $-0.35\sigma+\mu$, $-0.25\sigma+\mu$, $-0.11\sigma+\mu$ (if the first response was negative), where μ is the population mean WTP and σ is the population standard deviation assumed for the simulations. The follow-up bid values were calculated as the median

When ρ is equal to one, the appropriate statistical model is the interval-data model. In this case, however, a bivariate model would be overparametrized, but not incorrect, since the interval-data probit model is a special case of the bivariate probit model for correlation coefficient ρ tending to one.

When ρ is less than one (ρ=0.2, 0.6, 0.9 in the present simulation study) the correct statistical model is the bivariate probit. The interval-data estimates can be expected to be biased if the individual responses are generated by a bivariate WTP with correlation coefficient strictly less than one. In order to grasp why the estimates are likely to be biased, consider the contribution to the likelihood of each of the possible pairs of responses predicted by the two alternative models and reported in Table I. If the latent WTP's are jointly normally distributed, the correct probabilities of observing each pair of responses are read in column (2) of Table I. Assume now that an interval-data probit model is fitted, although the true model is a bivariate probit. If the true values of the parameters are plugged in the expressions for the contributions to the log likelihood of the interval-data probit model (column (1) of Table I), the probability of a ("no", "no") pair is overestimated, those of a ("no", "yes") and a ("yes", "no") pair are understimated, and that of a ("yes", "yes") pair is overestimated, since $\Phi(z_{i1}) \geq \Phi(z_{i1}, z_{i2}, \rho)$ for $\rho \geq 0$. The ML estimates from the interval-data probit model, $\hat{\theta}$, will adjust towards the values θ^*, possibly different from the true parameters θ, that make the counts of responses predicted by the interval-data model as close as possible to the actual counts. The bias of the interval-data estimates that does not vanish as the sample size increases, $\theta^* - \theta$, is likely to be small when the interval-data model constitutes a small departure from the bivariate probit (*i.e.*, for ρ close to one), and more substantial when the interval-data model is farther from the true model (*i.e.*, for low ρ).

The simulations were organized into two sets. No regressors were used to generate the latent WTP data in the first set of simulations, whereas a constant and two regressors, also drawn from a normal distribution, were included in the second set of simulations. In both sets for simplicity the data were

---

of the normal distribution truncated from below at the first bid value if the first response was positive, and from above at the first bid value if the first response was negative. A random subsample of 100 observations were assigned to each of the initial bids.

generated holding the parameters β and σ of the bivariate normal WTP constant across the two WTP equations, and the estimated bivariate probit models also imposed these restrictions. 500 replications were done using a sample size of 1000. Only the results of the simulations without covariates are reported and discussed in this paper.

Table II reports summary statistics of the simulations regarding the WTP model without covariates. Clearly, when the interval-data model is the correct statistical framework (when ρ is equal to 1), the interval-data estimates are virtually unbiased (the biases are 0.1% and 0.2%, respectively, of the true values of the parameters) and approximately normally distributed.[5] Table II also reports the empirical size for each of the coefficient, the nominal size being 0.05.[6] The empirical size for the intercept is close to 5%, whereas the test is slightly conservative for the scale parameter.

Surprisingly, the estimates of β_0 incur no substantial distortions as the correlation departs from one (in which case a bivariate probit would be the correct model): even for ρ as low as 0.2 the bias of $\hat{\beta}_0$ is no larger than -0.3% of the true value of β_0. This is also confirmed by the size for that coefficient, which is very close to 5% for all of the values of the correlation here considered.

[5]The bias of the estimate of a parameter θ is defined as $E(\hat{\theta}-\theta)$ and is easily computed from Table II as the difference between the average value of $\hat{\theta}$ over the replications and the true value of the parameter, θ.
[6]The empirical size is essentially the relative frequency of the replications in which the null hypothesis that the estimate of a parameter, $\hat{\theta}$, is equal to the true value of the parameter, θ, is rejected at the 5% significance level. The model and the estimates are well behaved when the empirical size of a parameter is approximately 5%. Biased estimates or other problems with the distributions of convergence of the estimates typically result in empirical sizes that are higher than the nominal size.

Table II. Simulation Results. True Model: Bivariate Normal with $\beta_0=2$, $\sigma^2=10$ and correlation coefficient ρ. 500 replications. Sample size: 1000.

Fitted Model: Interval-Data Model				
Intercept β_0	$\rho=0.2$	$\rho=0.6$	$\rho=0.9$	$\rho=1.0$ (a)
mean	1.9952	1.9973	1.9931	1.9971
standard deviation	0.0816	0.0933	0.1055	0.1081
min.	1.7062	1.6189	1.5801	1.6126
max.	2.2306	2.2379	2.2932	2.2830
skewness	-0.0514	-0.1066	-0.2203	-0.1023
kurtosis	3.0269	3.2747	3.4894	3.1042
empirical size (b)	0.0469	0.0420	0.0480	0.0460
scale σ				
mean	2.3595	2.7280	3.0894	3.1549
standard deviation	0.0709	0.0855	0.1032	0.1080
min.	2.1446	2.4454	2.7777	2.8261
max.	2.6087	2.9690	3.4072	3.4715
skewness	-0.0809	-0.0761	0.0083	0.0196
kurtosis	3.0686	3.3307	2.9738	2.6411
empirical size	1.0000	0.9980	0.1080	0.0380
Fitted Model: Bivariate Probit Model				
Intercept β_0	$\rho=0.2$	$\rho=0.6$	$\rho=0.9$	$\rho=1.0$ (a)
mean	1.9932	1.9965	1.9928	1.9972
standard deviation	0.0995	0.1046	0.1083	0.1094
min.	1.6362	1.5695	1.5677	1.5993
max.	2.2517	2.7212	2.2967	2.3010
skewness	-0.0593	-0.1201	-0.2267	-0.1044
kurtosis	2.9423	3.2714	3.5354	3.1365
empirical size	0.0520	0.0440	0.0560	0.0720
scale σ				
mean	3.1584	3.1610	3.1811	3.1928
standard deviation	0.2041	0.1983	0.1474	0.1261
min.	2.5533	2.6011	2.7999	2.8261
max.	3.8366	3.8223	3.8097	3.7389
skewness	0.1285	0.2137	0.5037	0.3503
kurtosis	3.3554	3.3122	3.6641	3.6032
empirical size	0.0540	0.0600	0.0440	0.0580
correlation ρ				
mean	0.2068	0.6092	0.9025	0.9528
standard deviation	0.0982	0.0916	0.0778	0.0571
min.	-0.0722	0.3453	0.6580	0.7323
max.	0.4870	0.9956	0.9999	0.9999
skewness	0.2972	0.1276	-0.5935	-1.3391
kurtosis	3.2133	3.5495	2.7506	4.3707
empirical size	0.0600	0.0620	0.0340	0.0640

(a) The interval-data model is the correct model when $\rho=1$. (b) The nominal size is 5%.

The interval-data estimate of the scale parameter, σ, is much more susceptible of becoming biased when the value of ρ is different from one. Fitting an interval-data model instead of the correct bivariate probit when ρ=0.2, 0.6 or 0.9 results in *downward* biases of the scale parameter, and in smaller variances of $\hat{\sigma}$ than with the bivariate probit. The bias of σ is very small when ρ=0.9 (about -2%) and results in an empirical size that is slightly in excess of the nominal size, but not by much (10%). The bias of σ becomes more pronounced for larger departures of ρ from one. When the correlation coefficient is 0.6, the average $\hat{\sigma}$ over the replications is 2.7280, a lower value than the true scale by about 14%, and even the largest estimate of σ over the replications (2.9690) is smaller than the true σ. The empirical size for $\hat{\sigma}$ is 99.98%, which means that $\hat{\sigma}$ is virtually always biased. For ρ=0.2 the average value of σ over the replications is about 2.35 (with a -25% bias), and the largest estimate was 2.6087, a value that is smaller than the true σ (3.1629).

In contrast, when the true correlation coefficient is 0.2, 0.6 and 0.9 and the correct statistical model, the bivariate probit, is fitted, the estimates for the coefficients are always effectively unbiased (the biases were no larger than -0.3% for β_0 and 1% for σ) and the empirical size are always very close to the 5% nominal size, with the only exception of the slightly conservative test for $\hat{\rho}$ in the simulation with true ρ equal to 0.6.

The *variances* of the estimates from the bivariate probit model are always larger than those of the interval-data estimates. The loss in efficiency for β_0 occurring when moving from an interval-data specification to a bivariate probit specification is very small when ρ=1.0, is still quite small but slightly larger when ρ=0.9, and becomes quite pronounced for ρ=0.6 and ρ=0.2. Because of their lower variances and negligible biases, the interval-data estimates of β_0 are superior in terms of the *mean square error* (MSE) criterion, even in the situations in which the interval-data model is the wrong model (*i.e.* when the correlation is strictly less than one).[7] The bivariate probit estimates of σ are markedly less efficient than

[7]The mean square error is defined as $MSE(\hat{\theta}) = Var(\hat{\theta}) + [Bias(\hat{\theta})]^2$ and is easily calculated from Tables II and III by adding together the square standard deviation of the estimates and the square difference of the average value of $\hat{\theta}$ and the true value of the parameter, θ. The MSE's for $\hat{\beta}_0$ are 0.0067 (ρ=0.2),

the interval-data $\hat{\sigma}$: their variances are at least about 1.4 times those from the interval-data model (for the case where ρ=1.0), become twice as large for ρ=0.9 and eight times as large for the lowest value of the correlation considered in this paper. In terms of the MSE criterion, the interval-data estimates of σ are better when the correlation is 0.9 and 1.0 (the MSE's being 0.0160 and 0.0117 vis-a-vis 0.0221 and 0.0168, respectively, for the bivariate probit model), but become inferior for smaller values of the correlation because of the large biases brought upon the interval-data estimates in those situations.

We were particularly interested in the behavior of $\hat{\rho}$ when there is perfect correlation between the WTP amounts. The average correlation coefficient over the replications was 0.9528, with individual values as low as 0.7323. The empirical size for ρ was only slightly in excess of the nominal size, a rather surprising result in light of our anticipation of more serious problems with the actual size when the true value of the parameter lies on the frontier of the parameter space. However, the histogram of the estimated ρ's (not reported) and the heavy skewness and kurtosis show that $\hat{\rho}$ is *not* normally distributed.

4. EFFICIENCY V. BIAS: EXAMPLES AND ALTERNATIVE MODELS.

In order to further explore the tradeoff of efficiency and bias, bivariate probit and interval-data models were fitted to the data collected through two well-known CV surveys. In the first example the empirical evidence supports the notion that the interval-data model might be preferred over the bivariate model because the difference in the WTP estimates is very small and the gain in efficiency from using the interval-data model quite large. The second example produces a larger difference between the estimates from the two alternative models. It is speculated that such difference might be due to the different degree of sensitivity of the alternative models to influential observations from the tails of the WTP distribution, and that other models should be devised that better suit the data.

0.0087 (ρ=0.6), 0.0112 (ρ=0.9) and 0.0117 (ρ=1.0) with the interval-data model, and (in order) 0.0100, 0.0109, 0.0118 and 0.0120 with the bivariate probit model.

Kakadu Conservation Zone CV Survey.

A contingent valuation survey was conducted by the Australian Resource Assessment Commission in 1990 as part of a cost-benefit analysis effort to evaluate options for the use of the resources of the Kakadu Conservation Zone (KCZ) (see Carson, Wilks and Imber [6]). The survey had a dichotomous-choice format with a follow-up. The survey instruments created two scenarios. The first depicted a more moderate ("MINOR") environmental impact on KCZ from mining, which was effectively limited to the areas of KCZ in the immediate vicinity of proposed mine sites, whereas the other ("MAJOR") implied a much higher environmental risk. To obtain a clean dataset, respondents that reported implausible WTP values were dropped from the sample following the method for consistency checks suggested in Carson, Wilks and Imber [6]. Cleaning the dataset did not alter the structure of the division of the minor-impact respondents among the categories of the answers, but did "trim" the upper tail of WTP somewhat by dropping some respondents that were assigned the highest bid values and had given ("yes", "yes") responses (see Appendix).

Only the results for the minor-impact subsample for untransformed WTP and without covariates are reported in this paper for ease of comparison with the Cameron and Quiggin paper.[8] The estimates for the mean/median WTP based on the sample used in this paper are just a little bit lower than those of the Cameron-Quiggin paper, as a result of excluding implausibly large values through the consistency checks and including respondents from the Northern Territories, who were found to have, on average, significantly lower values than the residents of the rest of Australia (Carson, Wilks and Imber [6]). There is virtually perfect agreement between the estimates of the correlation coefficient between the WTP variables in this paper and in the Cameron-Quiggin paper. The bivariate probit model of Table III assumes that $\mu_1 \equiv \mu_2$ and $\sigma_1 \equiv \sigma_2$ and obtains an estimated median WTP of A\$ 87.16, and a standard

---

[8]The assumption of normal WTP's may be inadequate. Carson, Wilks and Imber [6], for instance, fit a Weibull double-bound model to the survey data. The sample used in this paper is slightly larger than that used by Cameron and Quiggin because it came from a broader overall sample. Its composition is also likely to be somewhat different, because the present sample may have excluded some respondents of the Cameron-Quiggin sample that did not meet the consistency checks.

deviation of A$ 453.93. The value of the correlation coefficient is very close to one, being about 0.97. The interval-data estimate for median WTP, A$ 76.25, is reasonably close to that of the bivariate probit (A$ 87.16) and in fact the latter falls within the 95% confidence interval around μ based on the interval-data estimate (A$ 60.11-92.39). Furthermore, the standard error around $\hat{\mu}$ is 8.23, and is smaller than the standard error around $\hat{\mu}$ from the bivariate model (12.53) by about one-third. The variance of the interval-data estimate is thus *half* of that from the bivariate probit model. The length of the 95% confidence interval around μ is A$ 32.28 with the interval-data estimates, and A$ 48.54 with the bivariate probit estimates. The estimate for the standard deviation of WTP, σ, from the interval-data model is $265.18 and is less than that from the bivariate probit model, once again confirming the results of the simulation exercise.[9]

In this example the tradeoff between bias and efficiency appears to be in favor of using the interval-data model: the bias of the estimate of the mean/median WTP, if any, is negligible, and a large gain in efficiency over the estimates from the bivariate model is observed. These results can be contrasted with those of the Cameron-Quiggin paper to suggest that the large differences there detected between the sets of estimates from the alternative models might have been determined by an excessively thick upper tail of the WTP distribution.

Alaska CV Survey.

A CV survey was conducted in 1992 to measure the passive loss of benefits caused by the 1989 *Exxon-Valdez* oil spill in Prince William Sound (Carson et al. [5]). The breakdown of the respondents among the categories of answers was substantially different from that observed with the KCZ survey. The KCZ study appears to have the majority of the respondents concentrated in the ("no", "no") and ("yes",

---

[9]The value of the log likelihood is lower for the interval-data model than for the bivariate probit model, but we cannot use the standard likelihood ratio test to discriminate between these alternative models because, as discussed earlier, the distribution of this test statistic is not known when the restriction being tested (whether $\rho=1$) is on the frontier of the parameter space.

"yes") categories (approximately 40% and 46%, respectively, for the minor-impact subsample, and 35% and 50% for the entire sample), with very few falling in the other two classes. In the Alaska survey the bulk of the respondents appears to be in the ("no", "no") category (about 40% overall), with a marked increase in the ("yes", "no") class (about 25%) and a substantial decrease in the ("yes", "yes") category relative to the KCZ study. A surprising 30% of the Alaska respondents queried at the lower bid levels declined paying even as little as $5. Furthermore, there is no significant change in the distribution between categories of responses whether the initial payment question is $30 or $60, whereas the fraction of respondents whose WTP values fall between 10 and 30, 30 and 60, and 60 and 120 appears to change systematically with the questionnaire version. Overall, the incidence of "no" responses to the second payment question (64.43%) appears to be very high.[10]

The interval-data probit produces an estimated mean/median WTP of $ 45.51 and an estimated standard deviation of $ 109.42, which is, once again, much lower than any of the $\hat{\sigma}$ from the bivariate probit models. This time the 95% interval around the median from the interval-data estimate ($ 37.71-53.31) does *not* cover the bivariate probit median WTP. The standard error around the interval-data estimate is by about one-third lower than that around the bivariate probit median WTP from the first equation.

[10]Carson at al. [9] are aware of the large count of "no" responses to the follow-up question and offer an interpretation of this phenomenon based on the focus groups that preceded the survey. They suggest that respondents who answered "yes" to the first payment question are likely to say "no" in the follow-up because they think the goverment would waste the extra money requested to implement the plan to prevent oil spills described in the survey. Respondents who answered "no" to the initial payment question are also likely to answer "no" to the second payment question because they feel that the government would provide a good of inferior quality at the lower price, or that the government has a lower probability of providing the good at the lower cost amount.

Table III. Bivariate probit and Interval-data models of the Kakadu and Alaska CV surveys. (Standard errors in Parentheses).

	Australia's Kakadu CV Survey (Cleaned Data). Latent variable: WTP.		Alaska CV Survey Latent variable: WTP		Alaska CV Survey Latent variable: logWTP	
	Bivariate Probit	Interval-data Model	Bivariate Probit	Interval-data Model	Bivariate Probit	Interval-data Model
μ	87.16 (17.22)	76.25 (8.23)	28.78 (7.38)	45.51 (3.90)	3.0754 (0.1467)	3.366 (0.0711)
σ	453.93 (80.02)	265.18 (13.80)	191.50 (21.70)	109.42 (4.66)	3.6121 (0.4988)	2.003 (0.0846)
ρ	0.9694 (0.0103)	1	0.6975 (0.04)	1	0.6847 (0.0477)	1
sample size	1088	1088	1043	1043	1043	1043
log L	-1154.10	-1179.61	-1307.58	-1397.83	-1347.837	-1373.339
Implied Median WTP	A$ 87.16 (17.22)	A$ 76.25 (8.23)	$ 28.78 (7.38)	$ 45.51 (3.90)	$ 21.66 (3.1760)	$ 28.97 (2.0689)

The discrepancy between the WTP estimates could be the result of WTP being poorly approximated by a normal distribution, a finding also reported in Carson et al. [5]. We thus fitted a interval-data and bivariate probit based on the assumption of a *log* normal WTP. The interval-data model once again gives a lower $\hat{\sigma}$ than the value of the bivariate model and pegs the median WTP at $ 28.97. The bivariate model produces an estimated median of $ 21.66, which is not covered by the 95% confidence interval around the median from the interval-data model. The standard error around the interval-data estimate of median WTP is two-thirds that from the bivariate probit model. It is also found that the *mean* WTP is estimated at the seemingly reasonable value of $ 215.38 using the interval-data model, but jumps to a value of $ 14,750.51 with the bivariate probit model.

Choosing between a bivariate probit and an interval-data model would be quite difficult in this situation. Indeed, the discrepancy between the sets of estimates from the two models, the lower correlation coefficient between WTP's, and observation of the distribution among response categories (see Appendix) suggests that WTP is not very well explained as a normal (or as a log normal) or by the maintained mapping from latent WTP's to observed responses. Many have in fact suggested that moving to a mixture of models or a model with a "spike" (Johansson et al. [8]) at a WTP value of zero might be a sensible modeling strategy for the Alaska data.

5. CONCLUSIONS.

This paper explores the consequences of fitting an interval-data model to discrete choice CV data with a follow-up when the WTP responses are generated by a bivariate process, and fitting a bivariate model when the most parsimonious statistical framework is an interval-data model. The simulations of this paper show that the interval-data estimates for the mean/median WTP can be surprisingly robust to low values of ρ, although the interval-data model is technically wrong. As a result, an interval-data model is often superior to a bivariate model in terms of the criterion of the mean square error of the WTP estimates. This suggests that the researcher's model selection strategy might be in favor of the interval-data model, even though the correlation between WTP values is less than perfect, if the biases of the interval-data WTP estimates are sufficiently small relative to the gain in efficiency over the bivariate probit estimates. This is likely to be the case in many CV surveys in which the coefficient of correlation between a respondent's two WTP values can be expected to be close to one.

This paper further suggests that it is possible that the large biases of the estimates of the interval-data model found by Cameron and Quiggin [3] might be the result of, or might have been made worse by, a poor choice of the distribution assumed by the fitted interval-data model. The performances of the bivariate and the interval-data models in the presence of departures of WTP from normality are not well known, and therefore all results should be interpreted with caution. This point and the efficiency v. bias tradeoff are illustrated using two examples based on actual CV survey data. The first dataset is a "cleaned" version of the dataset used by Cameron and Quiggin, which appears to have effectively "trimmed" the upper tail of the WTP distribution. Because the interval-data estimates of mean/median

WTP are found to be very close to those from the bivariate model, and have a standard error that is smaller by about one-third, the interval-data estimates appear to offer a desirable tradeoff between bias and efficiency. In the second example the discrepancy between the bivariate probit estimates and the interval-data estimates of WTP hints at the need for a better model than the standard bivariate or interval-data equations, and is indeed a useful, informal diagnostic check of the assumptions about WTP usually made by CV researchers.

The lesson from these examples (and from the Cameron-Quiggin paper) is that bivariate probit and interval-data models are likely to provide different degrees of robustness to the presence of influential observations from the tails of the distribution of WTP or small departures from normality. Whether interval-data or bivariate probit models perform better appears to depend on the specific response patterns in the data under investigation. The bivariate probit model performed better than the interval-data model in the Cameron-Quiggin paper in terms of biases of the estimates, but a "trimmed" version of the same dataset produced WTP estimates that were very close, and markedly more efficient with the interval-data specification. The data from the Alaska CV survey points at the possibility that the bivariate model might be quite sensitive to the presence of responses "contaminated" by strategic considerations or "yea-" or "nay-saying" (Mitchell and Carson [10]). CV practitioners are advised to estimate both interval-data and bivariate probit models as part of the initial specification searches for the WTP model. They are also warned not to use the outcome of a test of the hypothesis that $\rho=1$ to choose between these alternative models, because this test is not well behaved when the correlation between the two WTP amounts is perfect.

APPENDIX

Table A.1. Summary of the Australia Kakadu Conservation Zone CV survey. Minor impact only.

First bid (Second bids)	% (no,no)	%(no,yes)	%(yes,no)	%(yes,yes)	N
5 (20, 2)	35.15	2.39	6.48	55.97	293
20 (50, 5)	35.76	3.82	5.56	54.86	288
50 (100, 20)	43.14	7.06	10.98	38.82	255
100 (250, 50)	49.60	7.54	8.33	34.52	252
Overall	**40.53**	**5.06**	**7.72**	**46.49**	**1088**
Overall in the Cameron-Quiggin paper	**35.20**	**4.94**	**8.98**	**50.84**	**1013**

Table A.2. WTP estimates from the Cameron-Quiggin paper. (Standard Errors in parentheses)

	Bivariate probit model	Interval-data model
Median/mean WTP (in A$)	150.22 (23.81)	510.10 (58.36)
Standard deviation of WTP (σ)	472.42 (80.34)	1437.81 (113.30)
Correlation coefficient ρ	0.9520 (0.01)	1
Log likelihood function	-1083.83	-1364.55

Table A.3. Summary of the Alaska CV survey data.

First bid (Second bids)	% (no,no)	%(no,yes)	%(yes,no)	%(yes,yes)	N
10 (30,5)	29.55	3.03	23.35	45.08	264
30 (60, 10)	36.70	11.61	25.84	25.84	267
60 (120, 30)	39.70	9.84	29.13	21.26	255
120 (250, 60)	54.09	11.67	20.62	13.62	257
Overall	**39.79**	**9.11**	**24.64**	**26.00**	**1043**

REFERENCES

[1] T. Amemiya, "Advanced Econometrics", Harvard University Press, Cambridge (1985).

[2] K. Arrow, R. Solow, P. R. Portney, E. E. Leamer, R. Radner and H. Schuman, Report of the NOAA Panel on Contingent Valuation: Natural Resource Damage Assessment Under the Oil Pollution Act of 1990, 58 *Federal Register* 4601 (1993).

[3] T.A. Cameron and J. Quiggin, Estimation Using Contingent Valuation Data from a "Dichotomous Choice with Follow-up" Questionnaire, *Journal of Environmental Economics and Management* (forthcoming).

[4] R.T. Carson, W. M. Hanemann and R. C. Mitchell, Determining the Demand for Public Goods by Simulating Referendums at Different Tax Prices, manuscript, University of California, San Diego (1986).

[5] R.T. Carson, R. C. Mitchell, W. M. Hanemann, R. J. Kopp, S. Presser, P. A. Ruud, A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill, Report to the General Attorney of the State of Alaska, (November 1992).

[6] R.T. Carson, L. Wilks and C. Imber, Valuing the Preservation of Australia's Kakadu Conservation Zone, *Oxford Economic Papers* (forthcoming).

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

[8] P.-O. Johansson, B. Kristom and H. Nyquist, Valuing a Public Good Using the CVM: Bid Vectors, Spikes, and Risk, unpublished manuscript, Stockholm School of Economics, Stockholm (1993).

[9] D. McFadden and Greg Leonard, Issues in the Contingent Valuation of Environmental Goods: Methodologies for Data Collection and Analysis, in J.A. Hausman (Ed.), *Contingent Valuation: A Critical Assessment*, North-Holland, Amsterdam (1993).

[10] R.C. Mitchell and R. T. Carson, "Using Surveys to Value Public Goods: The Contingent Valuation Method", Resources for the Future, Washington, DC (1989).

Three Stochastic Specifications of a Discrete/Continuous Choice Model of Demand Under Block Rate Pricing

Julie A. Hewitt*

* Department of Agricultural Economics and Economics, Montana State University.

I am indebted to Barbara Kanninen and John Antle for their instructive comments; also, I owe Michael Nieswiadomy the sincerest thanks for sharing his data with me.

Three Stochastic Specifications of a Discrete/Continuous Choice Model of Demand Under Block Rate Pricing

The residential demand for water services has been the subject of a lengthy literature. This literature has focused on the issues for both model specification and estimation introduced by the occurrence of block rate pricing of water services. This paper presents another view of specification of the demand model under block rate pricing, the discrete-continuous (D/C) choice model, and examines three stochastic specifications of the general model. The paper proceeds as follows. The first section presents a brief review of the issues in the water demand literature. In the second section, the D/C choice model is applied to behavior under block rate pricing, and in the third, the stochastic specification of the D/C choice model is addressed. In the fourth section, the results of estimating the D/C choice model, using a dataset that has appeared previously (Nieswiadomy and Molina, 1989), are presented and discussed. Summary and concluding remarks follow.

A BRIEF OVERVIEW OF THE LITERATURE

The main objective of most empirical studies of residential water demand is to calculate the price elasticity of demand. Unlike other demand literatures, however, very little of the debate in the water literature has focused on functional specification: the functional forms used are linear, semilogarithmic, power, and multiplicative forms (which include some power terms and some exponential terms). Estimation technique has been an issue for debate, though the discussion has been restricted to regression methods, including OLS, instrumental variables, and two- and three-stage least squares. A key observation to be made regarding water demand studies is that until Jones and Morris (1984), no study used disaggregate household level data; earlier studies used citywide or service area data which are assumed to describe the average household, but gloss over the heterogeneity of individual households.[1]

Regarding model specification and estimation technique, the literature may be summarized as follows. From Howe and Linaweaver's (1967) study forward, a persuasive argument existed in the literature to use marginal price in specifying a demand equation, rather than average price. After the work in the electricity demand literature of Taylor (1975) and Nordin (1976), the marginal price specification was modified to include a *difference* variable to account for the lump sum transfers implied by block rates.[2] The initial studies which followed then focused on the question of which model was the correct model: a specification using average price, or a specification based upon marginal price and *difference*. Later studies focused on the marginal price and *difference* specification (despite mixed evidence of pitting the two models) but used instrumental variable and two- and three-stage least squares estimation techniques to deal with the apparent simultaneity of price, quantity and *difference*. Price elasticity estimates in these studies range from zero to -1.57 in OLS studies (though OLS is biased when price is not independent of error term) and zero to -0.86 in simultaneous equations models with the exception of Deller, Chicoine and Ramamurthy's (1986) estimate of -1.12.

Nearly all of these studies mention in some fashion that while they model consumption directly, what is left unmodeled is the choice of block in which to locate consumption.

When water is sold according to a block rate, a serious issue for model specification and estimation which must be addressed is the obvious co-determination of the price, quantity and *difference*. A similar issue of co-determination is raised in the labor supply literature. The wage rate is a determinant of hours of labor supplied, yet the actual hours worked is a determinant of the true wage

[1]The three studies authored by (the group) Chicoine, Deller and Ramamurthy and the two by Nieswiadomy and Molina are the only empirical studies of water demand under block rate pricing which use the disaggregate data appropriate to a model of individual behavior in a billing period time frame.

[2]*Difference* is defined algebraically in the next section, but is equal to the difference in the cost of water if all units had the same marginal price and the cost according to the block rate. The motivation for including *difference* is that there must be some means of accounting for the fact that the marginal price is not necessarily the price of every consumed unit in a block rate situation.

rate, because of either variable marginal income tax rates or variable implicit tax rates for households participating in transfer programs, such as food stamps. As in the block rate pricing problem, the result is a budget constraint which is piecewise-linear. In the labor supply literature, Hall (1973) first recognized that the piecewise-linear constraint introduced issues of model specification and estimation. The first study that directly addresses the co-determination is that of Burtless and Hausman (1978). The literature is nicely summarized in Moffitt (1986), though the next section presents a brief summary of the model developed in the labor supply literature.[3]

THE DISCRETE/CONTINUOUS CHOICE MODEL: MICROECONOMIC THEORY

To derive the demand for a good which is priced according to a block rate schedule, we return to first principles, maximizing utility subject to a budget constraint, where the budget constraint of concern is piecewise-linear. We use the notation x for a vector of goods, with the first (or only) subscript the index of the good. Double subscripting of x denotes the block boundaries measured in the same units as the good, with x_{ij} denoting both the upper bound of the j^{th} block and the lower bound of the $j+1^{st}$ block. Prices are similarly subscripted, with p_{ij} denoting simply the marginal price for the j^{th} block of good i. Income is denoted by y; any fixed charges (such as a monthly service charge) are denoted by FC. We assume that only water is priced according to a block rate, and water is represented by x_1. With this notation, the full payment function for water consumption, $c(x_1)$ can be written as:

$$c(x_1) = \begin{cases} 0 & \text{if } x_1 = 0 \\ p_{11}x_1 + FC & \text{if } 0 < x_1 \le x_{11} \\ p_{12}(x_1 - x_{11}) + p_{11}x_{11} + FC & \text{if } x_{11} < x_1 \le x_{12} \\ p_{13}(x_1 - x_{12}) + p_{12}(x_{12} - x_{11}) + p_{11}x_{11} + FC & \text{if } x_{12} < x_1, \quad \text{etc.} \end{cases} \tag{1}$$

for a three block rate. If $p_{11}<p_{12}<p_{13}$, then the block rate is said to be increasing block (IB); if $p_{11}>p_{12}>p_{13}$, then the block rate is said to be decreasing block (DB); nonmonotonic rates are possible, as

[3]A more detailed development can be found in Hewitt and Hanemann (1994).

would be more complex assessments of fixed charges. The monotonic versions of equation (1) are the most ubiquitous form of block rate schedule.

To construct the budget constraint the consumer faces, we add to $c(x_1)$ the expenditures on all other (n-1) goods and services. Then we constrain this total expenditure to be less than or equal to income, noting that there are as many distinct expressions in the resulting budget constraint as there are in $c(x_1)$. The budget constraint is a piecewise-linear budget constraint, which can be characterized as a set of linear budget constraints, each relevant to only a particular range of consumption of x_1. As each distinct expression in $c(x_1)$, which we denote as $c^k(x_1)$, has a $p_{1k}x_1$ term, we rearrange terms to produce a budget constraint which more closely resembles the standard linear budget constraint. We do so by first defining the variable called *difference* pertinent to the k^{th} block as $d_k = p_{1k}x_1 - c^k(x_1)$, which, given (1), implies:

$$d_k = -FC - \sum_{j=1}^{k-1}(p_{1j} - p_{1j+1})x_{1j} \tag{2}$$

The *difference* can take on either negative or positive values, depending on whether the rate schedule is IB or DB, the block under consideration, *and* the magnitude of the service charge. If *difference* is positive, it acts as a subsidy or addition to income; if negative, *difference* is a lump sum tariff or reduction in income.[4] Adding *difference* to both sides of the budget constraint constructed from (1) gives the piecewise-linear budget constraint for the k^{th} segment:

$$p_{1k}x_1 + \sum_{j=2}^{n}p_jx_j \le y + d_k \qquad \text{for } x_{1k-1} < x_1 \le x_{1k} \tag{3}$$

With the set of budget constraints as in (3), we should find the solution to the utility maximization problem (the Marshallian demand functions) as follows. Maximize utility subject to the budget constraint implied by (and including) the $0<x_1\le x_{11}$ range constraint, then maximize utility subject to the budget constraint implied by the second block including the range constraint, and so on. These

[4]The *difference* terminology is due to Nordin (1976). Note that the definition used here is the negative of that suggested by Nordin and most commonly used. Notice also that d_k is a function solely of the exogenous payment function or rate structure, and not dependent upon any endogenous variables.

solutions are the continuous choices, and are conditional upon the quantity meeting the range constraint for that budget segment. The overall maximum is determined by comparing these conditional choices to see which yields the highest (indirect) utility, which is the determination of the discrete choice. A particular level of consumption is then best described as a combination of the discrete and conditional continuous choices.

Figure 1 shows a two-block increasing-block rate situation, where the utility maximization bundle includes consumption of x_1 in block two. In the figure, the continuous choice conditional on block one is the upper boundary of the range of block one (with the implied consumption of all other goods). That is, given the range and budget constraints for the subproblem of the first block, the highest utility attainable, while not the result of a tangency between the budget segment and an indifference curve, is nonetheless readily identifiable as the upper block boundary.[5] Furthermore, this point is also part of the utility maximization subproblem of the second block and an interior solution in the second block (given convexity of indifference curves) will necessarily be associated with a higher utility level than the lower block boundary.[6] Figure 1 shows indifference curve U^1 as the highest attainable utility level given the first budget subset constraint, and U^2 as that for the second budget subset. U^2 in this case is also the unconditional maximum, making the second block the discrete choice. A further possibility (not shown) is that the highest attainable indifference curve is not tangent to any

[5]Moffitt (1986) implied that the form of the direct utility must be known to make the choice between a conditional continuous choice which is an interior point (using the indirect utility and actual prices, and income plus difference) and a conditional continuous choice which is a boundary point (using the direct utility because actual prices don't produce a tangency so indirect utility cannot be used), that is, to make the discrete choice. The utility index of a boundary point can be calculated from the indirect utility function so long as virtual prices and income are arguments. See Neary and Roberts (1980) and Rothbarth (1941).

[6]In the case of a decreasing block rate, two points should be noted. First, an interior solution to one utility maximization subproblem, given the convexity of indifference curves, does not rule out an interior solution to another utility maximization subproblem. These two solutions may be tangencies to the same indifference curve, which leaves the discrete choice (which block to locate in) ambiguous. Secondly, the possible double tangency implies a range of consumption that will never be observed by the utility maximizing household. Studies which have dealt with the issues raised by the decreasing block rate anomalies (including the need to numerically approximate the likelihood function) and to which the reader is referred include Burtless and Hausman (1978), and Reece and Zieschang (1985).

budget subset, but goes through the kink point, where $x_1 = x_{11}$ making the discrete choice the kink point itself.

An interesting feature of the D/C choice model that we think bears some emphasis is that for a given functional form of direct utility, all the conditional demands produced by maximizing utility subject to the budget constraints will have the same functional form as the utility maximization subject to standard linear budget constraints. They will, however, differ in their arguments, specifically in the marginal price and income (via the *difference* variable) associated with each block. The underlying parameters of the utility function whose values we seek in estimation are assumed to be the same regardless of discrete choice, that is, the household's tastes are not assumed to be functionally dependent on decisions.[7] Taking advantage of this feature, we denote the functional form of the standard Marshallian demand as $\bar{x}_1$, which will take on the appropriate arguments for each conditional demand expression. We can express the unconditional Marshallian demand for x_1 as the combination of discrete and conditional continuous choices now, assuming for simplicity a two-block increasing-block rate, as follows:

$$x_1 = \begin{cases} \bar{x}_1(p_{11}, y+d_1;\theta) & \text{if } \bar{x}_1(p_{11}, y+d_1;\theta) < x_{11} \\ x_{11} & \text{if } x_{11} \le \bar{x}_1(p_{11}, y+d_1;\theta) \text{ and } \bar{x}_1(p_{12}, y+d_2;\theta) \le x_{11} \\ \bar{x}_1(p_{12}, y+d_2;\theta) & \text{if } x_{11} < \bar{x}_1(p_{12}, y+d_2;\theta) \end{cases} \tag{4}$$

where θ denotes the underlying parameters of the utility function. Each line of the bracket portion of (4) specifies on the left the conditional continuous choice, and on the right the condition for the

[7]This is an assumption that could be relaxed, though we do not do so in this study. The relaxation of this assumption could also be due to assumed differences in the utility maximization problem when the block rate is increasing versus decreasing and would lead to separate estimation of subsets of data under increasing and decreasing rates, such as in Nieswiadomy and Molina (1989).

particular discrete choice upon which the continuous choice is conditioned. Equation (4) is not a stochastic model and therefore is not yet a model which can be statistically estimated.

THE D/C CHOICE MODEL: ECONOMETRICT SPECIFICATION

In the standard econometric specification of a deterministic economic model (one that is not otherwise based explicitly upon a theory of risk or uncertainty), a random error term that is normally (lognormally) distributed with mean zero (one) is added (multiplied), and justified as explaining the influence of unobserved or omitted explanatory variables. Since the D/C choice model starts out as a more complex statement of demand behavior, we must be cautious in how we introduce random error terms: there are several different ways in which random errors can be added, and each has a distinct interpretation. For simplicity, we continue in this section with the two-block, increasing-block rate as in (4). Three models are presented below; Moffitt (1986) surveys these three models (focusing on the most general version) as they appear in the literatures on labor supply, welfare programs and charitable contributions.

Suppose we add an error term, η, to each conditional demand, $\bar{x}_1$ in (4). What explanation could justify this specification? Assume the relationship $x_1 = \bar{x}_1() + \eta$ holds where $\bar{x}_1()$ is planned consumption, x_1 is observed consumption, and η is the difference between the two. This error term has been variously called measurement error (Burtless and Hausman, 1978) and optimization error (Pudney, 1989), but we find the term perception error to be more general (Hewitt, 1993).[8] Clearly, then, η is added to the conditional continuous choice expressions in (4), but η would also appear in the discrete choice expressions on the right. That is, because $\bar{x}_1()$ is not observed, it cannot be used alone in

[8]Measurement error may be a bit of a misnomer; measurement error implies that the error is due to the econometrician's imperfect measurement abilities including meter error. Optimization error may also be a bit of a misnomer; optimization error implies that the error is due to the cognitive abilities of the consumer or the inability of the consumer to attain the optimum point due to unforeseen forces. Hewitt (1993) uses the more general term of perception error, noting that it matters not which party the error is actually due to, for it cannot be attributed to either in the absence of additional information, and may in fact be due to both.

describing the discrete choice. Thus $x_1-\eta$ is substituted for $\bar{x}_1$ and the discrete choice expression for block one becomes $x_1-\eta<x_{11}$, which is equivalent to $\eta>x_1-x_{11}$. The stochastic form of (4) with perception error is:

$$x_1 = \begin{cases} \bar{x}_1(p_{11}, y+d_1;\theta) + \eta & \text{if } \eta > x_1 - x_{11} \\ x_{11} + \eta & \text{if } \eta = x_1 - x_{11} \\ \bar{x}_1(p_{12}, y+d_2;\theta) + \eta & \text{if } \eta < x_1 - x_{11} \end{cases} \tag{5}$$

Equation (5) is the basis of the likelihood of observing x_1 given values of the exogenous variables. If f is the pdf of η and F is the cdf of η, the probability of block one being the discrete choice is $\Pr(\eta>x_1-x_{11}) = 1-F(x_1-x_{11})$. The probability of block two being the discrete choice is similarly $F(x_1-x_{11})$.[9]

In constructing the likelihood of an observed value for x_1 from (5), we should note that the discrete choice in this situation is unobserved. That is, an observed value for x_1 does not tell us what block planned consumption, $\bar{x}_1()$ falls into, only what block observed consumption, $\bar{x}_1() + \eta$, is in. Unable to ascertain the error or planned discrete choice, the likelihood of an observation is the sum of the probabilities associated with each possible discrete and conditional continuous choice pair, as in:

$$\begin{aligned} \Pr(x_1) = {} & f_\eta\left(x_1 - \bar{x}_1(p_{11}, y + d_1;\theta)\right)\left[1 - F_\eta(x_1 - x_{11})\right] + \\ & + f_\eta\left(x_1 - \bar{x}_1(p_{12}, y+d_2;\theta)\right) F_\eta(x_1 - x_{11}) \end{aligned} \tag{6}$$

The source of the random term, η, is a variant of the source often cited in a regression model: η picks up the effect of unobserved or omitted variables *which explain the discrepancy between planned and actual consumption.* The parameters of this likelihood can still be identified, but there is no separation of the observations by discrete choice.

The effect of the lack of sample separation is apparent when noting that the log-likelihood of an observation is the logarithm of the sum in (6), rather than the more common sum of logarithms, which occurs when sample separation exists. The log-likelihood, its gradients and its Hessian are more

[9] The probability of the kink point being the discrete choice would appear at first glance to be $f(x_1-x_{11})$, but for continuous random variables, the probability density function evaluated at a particular value does not represent the probability that the random variable takes on that value. See, for example, Mood, Graybill and Boes (1974, p. 61).

complex than those of log-likelihoods with some degree of sample separation embedded, causing more computer resource intensive maximization. Furthermore, while this model dismisses the misinformation that occurs when actual consumption is used to indicate the planned discrete choice, it may go too far in dismissing discrete choice information. Surely for some observations the planned and actual consumption levels indicate the same discrete choice, yet such information is also dismissed.

For an alternative stochastic specification, suppose that the source of the error is unmeasurable heterogeneity in household preferences, and that this heterogeneity is embodied in some parameter of the vector θ.[10] Thus, heterogeneity appears in the $\bar{x}_1$ conditional continuous choice expressions, known to the household at the time of consumption decision-making, but unknown or only imperfectly visible to the econometrician. Let ε be the random term due to household heterogeneity which we now express separately from θ, additively as in $\bar{x}_1 + \varepsilon$. We then rewrite (4) as

$$x_1 = \begin{cases} \bar{x}_1(p_{11}, y+d_1; \theta) + \varepsilon & \text{if } \varepsilon < x_{11} - \bar{x}_1(p_{11}, y+d_1; \theta) \\ x_{11} & \text{if } x_{11} - \bar{x}_1(p_{11}, y+d_1; \theta) \le \varepsilon \le x_{11} - \bar{x}_1(p_{12}, y+d_2; \theta) \\ \bar{x}_1(p_{12}, y+d_2; \theta) + \varepsilon & \text{if } x_{11} - \bar{x}_1(p_{12}, y+d_2; \theta) < \varepsilon \end{cases} \tag{7}$$

There are a few differences between equations (5) and (7) to note. First, the conditional continuous choice where the discrete choice is the kink point is exactly x_{11}, not x_{11} plus an error term — and observations can occur at the kink with nonzero probability mass. Secondly, unlike in equation (5) where just the conditional continuous choice contains $\bar{x}_1()$ terms, in (7) both the continuous and discrete choices are directly a function of the conditional continuous demands. The probability of block one being the discrete choice is $\Pr(\varepsilon < x_{11} - \bar{x}_1(p_{11}, y+d_1; \theta)) = F(x_{11} - \bar{x}_1(p_{11}, y+d_1; \theta))$. Similarly, the probability that block two is the discrete choice is $1 - F(x_{11} - \bar{x}_1(p_{12}, y+d_2; \theta))$ and the probability of the kink point being chosen is: $F(x_{11} - \bar{x}_1(p_{12}, y+d_2; \theta)) - F(x_{11} - \bar{x}_1(p_{11}, y+d_1; \theta))$. Finally, there is

[10]Household heterogeneity could have an explained component that is a function of sociodemographic variables, possibly even of prices or expenditures. The element of θ which embodies heterogeneity could be a parameter of the utility function, or a scaling or translating factor. See Pollak and Wales (1981), Lewbel (1985), Pudney (1989, pp 34-39), and Hewitt (1993, Chapter 5) for more on introducing heterogeneity in a utility theoretic fashion.

sample separation: if x_1 is observed to be less than (greater/equal) x_{11}, then block one (two/kink) is the true discrete choice. Not only is the discrete choice observed without error, so too is the continuous choice.[11] Thus, the likelihood of an observation, x_1, is:

$$\Pr(x_1) = \begin{cases} f_\varepsilon\left(x_1 - \bar{x}_1(p_{11}, y + d_1)\right) & \text{if } x_1 < x_{11} \\ F_\varepsilon\left(x_{11} - \bar{x}_1(p_{12}, y + d_2)\right) - F_\varepsilon\left(x_{11} - \bar{x}_1(p_{11}, y + d_1)\right) & \text{if } x_1 = x_{11} \\ f_\varepsilon\left(x_{11} - \bar{x}_1(p_{12}, y + d_2)\right) & \text{if } x_1 > x_{11} \end{cases} \tag{8}$$

The source of the random term, ε, is a variant of the source often cited in a regression model: ε picks up the effect of unobserved or omitted variables *which explain the heterogeneity of households.*[12]

The household heterogeneity source of error would be a plausible explanation except for the following two points. First, unless the utility function parameters are such that marginal price and income effects are zero (i.e., that $\bar{x}_1(p_{11},y+d_1;\theta) \equiv \bar{x}_1(p_{12},y+d_2;\theta)\ \forall\ p_{11}, p_{12}, d_1, d_2$), a nonzero probability mass would occur at the kink point.[13] The nonzero probability mass at the kink implies that there would be a jump discontinuity in the frequency distribution of observed x_1. This jump discontinuity is rarely observed in actual data generated under a block rate pricing regime.

More importantly (and independent of whether a jump discontinuity in the distribution of x_1 can be discerned or not), it is highly unlikely that the discrete choice is observed without error. A household which would have chosen to locate in a particular block but near the kink would necessarily have a truncated distribution for their random error term in order for their consumption to actually remain in the same block. (We could call such a model an ordered tobit.) While the discrete choice probabilities in (7) account for the necessary truncation, it is simply unrealistic to assume that

---

[11]Though households are heterogeneous, each knows its planned consumption level and actually consumes that level so there is no meaningful distinction between the planned and actual. The source of the error term then is that the econometrician cannot perfectly explain the heterogeneity of household preferences.

[12]The log-likelihood is expressed as three summations of log-probabilities, each summation over one of three mutually exclusive and exhaustive sets of observations.

[13]When p_{11} and p_{12} are not equal and when the cdf F_ε is strictly monotonically increasing, the zero price and income effect condition is also necessary for there to be a nonzero probability mass at the kink.

households which choose to locate interior to a block but near the kink have a random error distribution different from households choosing interior points not near the kink point (hence less severely truncated). That is, households are masters of their own tastes and preferences but do not set the kink points, while the water utility sets the kink points but has little to say about household tastes and preferences. As θ and x_{1k} are independent, there is little reason for the econometrician to specify the heterogeneity of households in a way so critically dependent upon the x_{1k}.[14]

Keeping the criticisms of the two models presented thus far in mind, a third model which combines the two previous models is presented next. This model is due to Burtless and Hausman (1978). Heterogeneity error, ε, continues to drive the discrete choice because it is an error to the econometrician but not the household, but there is a second source of error, denoted η, which represents the difference between planned and actual consumption. Thus, equations (5) and (7) are combined, as in:

$$x_1 = \begin{cases} \bar{x}_1(p_{11},y+d_1;\theta)+\varepsilon+\eta & \text{if } \varepsilon < x_{11}-\bar{x}_1(p_{11},y+d_1;\theta) \\ x_{11}+\eta & \text{if } x_{11}-\bar{x}_1(p_{11},y+d_1;\theta) \le \varepsilon \le x_{11}-\bar{x}_1(p_{12},y+d_2;\theta) \\ \bar{x}_1(p_{12},y+d_2;\theta)+\varepsilon+\eta & \text{if } x_{11}-\bar{x}_1(p_{12},y+d_2;\theta) < \varepsilon \end{cases} \tag{9}$$

If ε and η are independent, then adding perception error would smooth out the jump discontinuity in the distribution of x_1 observations. The difference between the two-error version and the heterogeneity one-error version of the D/C choice model is shown in Figure 2.

Equation (9) is the basis for writing the likelihood of an observation. For the kink point solution, the continuous choice is based on a particular value of η and the discrete choice on a range of values for ε; thus we need to specify the joint distribution, $f_{\varepsilon,\eta}(\varepsilon,\eta)$. The continuous choice expression for the two interior solutions is based upon a particular value of $\varepsilon+\eta$, while the discrete choice is based upon ranges of values of ε. Thus to construct the probability statement for block one and block two interior points, we also need the joint distribution:

[14]This is not to suggest that θ and x_{1k} are always independent. Utilities which adopt public education or advertising campaigns in periods of water shortage surely affect θ, while households with certain values for θ may attempt to influence the utility's setting of x_{1k} as in Martin *et.al.* (1984).

$$g_{\varepsilon+\eta,\varepsilon}(\varepsilon+\eta,\varepsilon) = |J|\, f_{\varepsilon,\eta}(\varepsilon,\eta)\, I_g(\varepsilon,\eta) \tag{10}$$

where $|J|$ is the Jacobian of the transformation from (ε,η) to $(\varepsilon+\eta,\eta)$ and I is the indicator function for the range of g.[15]

This two-error model specification may also be described as a likelihood having no sample separation (by discrete choice). Any observation, regardless of which block it is *observed* ex poste to have chosen, could have been the result of any of the three discrete choices, in concert with a perception-error term of sufficient magnitude to place the observed value in a different discrete range. This adds richness to the model in the sense that not only do we directly address the discrete choice component of the model (as in the heterogeneity-error-only model), but (akin to the perception-error-only model) we do so without forcing the discrete choice to be exactly the *observed* ex poste choice. However, this richness comes only at the expense of the more complex likelihood, hence more complex gradients and Hessian, and therefore slower convergence to the set of parameters which maximize the likelihood.[16]

We turn now to the empirical application of the models presented.

THE DATA, ESTIMATION AND RESULTS

The data used in this study has been used previously (Nieswiadomy and Molina, 1988, 1989) and the reader is referred to these studies for more detailed descriptions of the data than that contained herein. The data used here is the same, with the exceptions noted below. Generally these differences were dictated by differences in the underlying economic models the data was used to estimate and do not represent improvements to the data that would otherwise make re-estimation of the Nieswiadomy

[15]Although the Jacobian for this transformation is simply one, equation (10) is given in a somewhat general form to highlight the fact that the additive error structure in (9) is part of the distributional assumption.

[16]Indeed, Newton and quasi-Newton search algorithms appear to have a smaller advantage in maximizing the likelihoods in (5), (7) and (9) over the other methods, such as grid search and steepest ascent. Part of the advantage of Newton methods could be retrieved by the use of judiciously chosen starting values and is an area for further research.

and Molina (N&M) models appropriate. Although the N&M dataset is a pooled cross-section, time-series of monthly observations of 121 households over ten years, this study uses only 1703 observations: those from June, July or August months with two block increasing block rates in effect.[17] Table 2 shows summary statistics of the subset of the data used. As in N&M, income is constructed from the value of the home for tax assessment purposes and, for this study, the monthly income is converted to a billing period equivalent, using the number of days in a billing period. Income is assumed to reflect real income, and hence is not deflated.[18] Prices, and, hence, the calculated *difference* variables are in nominal terms and are thus deflated. Lawn size is constructed as the lot size minus two times the house size.[19] Finally, the number of bathrooms and days in the billing period were used as variables in this study though they proved to be insignificant in N&M.

The use of the variable number of days per billing period bears some explanation. In their study N&M convert the dependent variable, x_1, to a 30-day equivalent for consistency. The water utility does not do the same, however, when determining a household's bill. That is, the likelihood that a household consumes in the second block is higher, *ceteris paribus*, the longer the billing period. Put another way, the utility does not adjust one's marginal rate because the household had a 45 day billing

---

[17]The subsetting of the data was done for two reasons. The primary reason was to reduce the number of observations to increase the speed of estimation of the maximum likelihood model. Secondly, the likelihood function for an IB rate is simpler than the likelihood for a DB rate, because there are terms that can only be numerically expressed in the DB likelihood (over and above the normal cdf term which is included in both likelihoods and is easily programmed). See Reece and Zieschang (1985) on this. Thirdly, the likelihood in (9) is simpler the fewer are the discrete choices, so we restrict ourselves to two blocks and the kink between them. The total number of usable two block IB observations is 5935. Finally, to the extent that possibly different processes determine winter and summer water consumption, we reduce the potential for misspecification by restricting the data to three summer months, or 1703 observations.

[18]Income as measured here essentially reflects permanent income rather than the sum of permanent and transitory income. We suggest that permanent income is more likely to be the determining factor for water consumption than transitory income fluctuations. The assessment data is from the year 1984 and is in 1984 dollars. Other variables expressed in dollars will be converted to 1982-1984 dollars using the CPI-U, all items, all urban consumers, indexed to 1982-1984 = 100 (U.S. Department of Commerce, 1989).

[19]For some observations, a one story house assumption led to negative lawn size, and so a two story assumption was maintained for all houses, in the absence of more detailed information on variation in structure.

period but would have consumed within the first block had the billing period been only 30 days. Thus we leave x_1 as registered on the meter and use days as an explanatory variable. The weather variable is exactly as constructed by Nieswiadomy and Molina, and measures in inches the potential evapotranspiration for Bermuda grass less actual rainfall, with both potential evapotranspiration and actual rainfall based on observed daily weather for the days comprising each household's billing period.

Preliminary analysis of the data indicated that the frequency distribution of the dependent variable was significantly skewed to the right, while little skewness existed in the independent variables. Use of a symmetric distribution in specifying a likelihood function was thought to work against accurate model estimation and so the lognormal distributional assumption is maintained.[20] A power form is chosen for demand, resulting in the expression of the conventional demand:

$$\bar{x}_1(p_1, y; Z\delta, \alpha, \mu) = \exp(Z\delta)\,(p_1)^{\alpha}\, y^{\mu} \tag{11}$$

The stochastic expression takes the form:

$$x_1 = \bar{x}_1 \exp(\varepsilon) \exp(\eta) \tag{12}$$

The likelihood is based upon:

$$\ln x_1 = \begin{cases} Z\delta + \alpha \ln p_{11} + \mu \ln(y + d_1) + \varepsilon + \eta & \text{if } \varepsilon < \ln x_{11} - Z\delta - \alpha \ln p_{11} - \mu \ln(y + d_1) \\ \ln x_{11} + \eta \quad \text{if } \ln x_{11} - Z\delta - \alpha \ln p_{11} - \mu \ln(y + d_1) \le \varepsilon \le \ln x_{11} - Z\delta - \alpha \ln p_{12} - \mu \ln(y + d_2) \\ Z\delta + \alpha \ln p_{12} + \mu \ln(y + d_2) + \varepsilon + \eta & \text{if } \varepsilon > \ln x_{11} - Z\delta - \alpha \ln p_{12} - \mu \ln(y + d_2) \end{cases} \tag{13}$$

where Z represents the sociodemographic variables that affect the constant of proportionality, ε represents heterogeneous preferences error, η represents perception error, and δ, α, and μ are the

[20]If ε, η are independently normal, then $\exp(\varepsilon)$, $\exp(\eta)$ are independently lognormal. The heterogeneous preferences error is hypothesized to be skewed for much the same reason that the distribution of the dependent variable is: the constant of proportionality is clearly going to have to explain much of the skewness in consumption, since prices and income are not skewed enough to be the full explanation. The perception error is hypothesized to be skewed for more concrete reasoning. Perception error includes the following sources of error: rounding error in recording consumption, meter error, and leaks. While rounding error would be uniform over the rounding interval, meter error and leaks would likely have distributions which are highly skewed to the right. The total of all the sources of perception error then is also skewed to the right.

unknown parameters of the utility function.[21] The full log-likelihood functions for the three stochastic specifications of the D/C choice model, using the distributional and functional assumptions of (11) to (13), are shown in the appendix.

According to the second order conditions of utility maximization, the sign of the p_1 exponent should be negative and the y+d exponent should be positive. Intuition, not economic theory per se, suggests that the signs of all the coefficients on the sociodemographic variables included in Z would be positive (weather reflects the water needs not met by rainfall). The sociodemographic variables affect demand multiplicatively through $\exp(Z\delta)$, hence, there is no theoretical restriction on the signs of the elements of δ.

The final a priori expectation is that $\sigma_\eta<\sigma_\varepsilon$, which is to say that more of the unexplained variability of observed consumption is due to the econometrician's inability to fully characterize heterogeneous preferences than to random impercipience, whether on the part of the econometrician or the consuming household. This expectation is based on the self-interest motivation of both the household and the water utility to take into account the factors which lead to impercipience in determining their behavior.

The results of four models are presented in Table 2. In addition to the three stochastic versions of the D/C choice model developed in this paper, we include the results of a OLS regression for comparison. The OLS regression is essentially the same functional and distributional form as the D/C choice models, with two caveats. First, the single error term captures either type of error or both; in an OLS model, the errors cannot be separately identified and we denote the standard error of the regression as the sum of σ_ε and σ_η. Secondly, the choice of marginal price and *difference* values to use for each observation is based upon the values appropriate to the block of observed consumption. The clearly

[21]Foster and Beattie (1979) suggest the use of an exponential form rather than a power form, at least with respect to the price variable specification, because the power form rules out eventual satiation in water consumption due to a constant elasticity. As will be shown below, the power form does not imply a constant elasticity in the D/C choice model.

implies that the error terms and independent variables are not independent; we show these results for comparison only.

The first thing we note about the results is that the coefficients on the sociodemographic variables are fairly consistent across models, at least in the sense of being on the same order of magnitude. Next, with one exception, the signs on the sociodemographic variables in all four models are positive as we expected (but don't require on grounds of theory). The exception is house size in the OLS and heterogeneity-only models. Generally, the constant and house size are insignificant while lawn size, weather, bathrooms, and days generally are significant. More importantly, with respect to theoretical considerations, income plus *difference* has the expected positive effect on consumption (though not in a very significant fashion), while marginal price is always significant, though it has the expected negative effect only in the two-error model. Next we note that the standard deviations in the two-error model have the expected relationship where $\sigma_\varepsilon > \sigma_\eta$, here more than doubly so.

We compare these models to each other now, though only in an ad hoc fashion.[22] Perhaps the most striking result in terms of overall model comparison is the similarity of the results for the OLS and heterogeneity-error-only models. The parameter estimates are identical for all intents and purposes, though the t-statistics differ. When we consider the fact that both models use the observed consumption level to indicate the appropriate values of marginal price and *difference*, this is less surprising; indeed, given that the budget constraint is kinked but not greatly so, that water expenditures are generally not a large portion of income, and that few observations occur at zero or the kink point, we might imagine the improvement of the (ordered) tobit estimates over the OLS estimates to be minor. The similar use of observed consumption to indicate values for marginal price and *difference* appears to be much more important in driving the parameter estimates.

[22]The validity of statistical tests of model comparison are being pursued currently. In the absence of such tests, it should be noted that the two one-error models **are not** nested within the two-error model. This nesting is in the usual sense that parameter restrictions on the general model produce the nested model and the statistical validity of the restrictions may be tested. The two-error likelihood cannot be evaluated either when $\sigma_\varepsilon = 0$ or $\sigma_\eta = 0$.

Given that the OLS and heterogeneity-error-only models produce such similar results and are based on the same key assumption (consumption is observed without error), how then do these models compare to the other two? As alluded to earlier, if the heterogeneity-error-only model takes the observed consumption level as a perfect indicator of true consumption, and hence the discrete choice, then the perception-error-only model takes the completely opposite approach by not using the observed consumption level at all as an indicator of the discrete choice. That is, if the information contained in the observed consumption level is taken completely at face value in the heterogeneity-error-only model, it is completely discounted in the perception-error-only model.

If truth is somewhere in between these two polar cases, that is, that there is some information as well as some misinformation regarding the discrete choice contained in the observed consumption level, then we might characterize the perception-error-only model as a limited information maximum likelihood model and the heterogeneity-error-only model could be characterized as a hyper-information maximum likelihood model. That both of these models violate the second-order conditions for utility maximization is not then terribly surprising.

The two-error model takes the approach that truth is somewhere in between, that there is both information and misinformation in the observed consumption level. For this reason, we characterize this version as a full information maximum likelihood model, and are heartened by the fact that this model satisfies the second-order conditions. Note that the asymptotic t-statistics for σ_ϵ and σ_η indicate that each is significantly different from zero, indicating that the two-error model is superior to either one-error version. We turn now to a discussion of the price elasticity of demand (the ultimate objective of most studies of residential water demand), focusing only on the two-error, full information maximum likelihood results.

Given the power form used for price and income in the expression of conditional demand, we might be tempted to conclude (based on the two-error model) that the price elasticity of the demand for water is constant at -1.90 and the income elasticity is constant at 0.18. However, these elasticities are the elasticities associated with the conditional demands only. If price and/or income change, so do the probabilities of making particular discrete choices and the conditional demands associated with each discrete choice. The method of calculating elasticities of the D/C choice model is as follows: equation (9) is rewritten to reflect our functional and distributional assumptions and the expectation is taken of this expression for x_1 from which we could either differentiate to calculate the elasticity or numerically integrate (because the expression contains conditional expectations and cdf evaluations).

The expectation of x_1 can be taken either over ε alone or over both ε and η. If the expectation is over ε only, then the expectation is that of planned consumption, or x_1 without the influence of η (random impercipience). The expectation over both ε and η is that of observed consumption, that is x_1 as influenced by economic variables, household heterogeneity, and random impercipience (the distribution of which we do have some information on). Table 3 shows sample and predicted values of consumption evaluated at both the mean and median values of the independent variables. The last column gives the values of the first block conditional demand for comparison.[23]

Table 4 shows various elasticity estimates from numerical integration. The price elasticity calculations assume that both p_{11} and p_{12} change in the same proportion (including their subsequent effect on the *difference* variable value). This is a somewhat arbitrary assumption, but nonetheless suits the purpose at hand. The income elasticities assume a change in income, or the negative of a change in a fixed charge, but no change that would otherwise cause the *difference* variable to change (e.g., changes in FC, p_{11} or x_{11}).

There are several points worth noting regarding this table. First, note that the elasticity of the expectation of observed consumption is always smaller (in absolute value) than the elasticity of the expectation of planned consumption, for both price and income. Next note that although the functional form assumed for the conditional demand is really a constant elasticity form for both price and income,

[23]This is due to the lognormal distributional assumption for both ε and η. As $\exp(\varepsilon)$ and $\exp(\eta)$ are normally distributed with mean zero and reported standard deviations, ε and η are lognormally distributed with $E[\varepsilon] = 1.4647$ and $E[\eta] = 1.0734$.

the elasticities do depend on the values of the explanatory variables at which they are evaluated.[24] As the elasticities for this functional form and the single-price one-error OLS model would be just the parameter estimates for price and income, the calculated elasticities for the block-rate two-error model are less elastic than would be suggested by interpreting the exponents on price and income as elasticities. However, the elasticities reported here are nearly uniformly more elastic than those reported in the literature. Interestingly, the most price elastic study in the literature is also one of the first; Howe and Linaweaver (1967) estimate a summer price elasticity of -1.57. Only one other study suggests that demand is price elastic: Deller, Chicoine and Ramamurthy (1986).

CONCLUSIONS AND FURTHER RESEARCH

In this paper, a discrete/continuous choice model of the residential demand for water under block rate pricing is presented in several forms and estimated. The empirical analysis uses a dataset of a previously published study of Nieswiadomy and Molina (1988, 1989). The power function and lognormal error form assumptions are similar to the four models estimated. The striking result is that the D/C choice model produces a price elasticity estimate of -1.13 to -1.59, which is much more elastic than all but two previously published estimates (based on models which may well be misspecified).

The model responds to a criticism of economic or econometric studies of the residential demand for water services often made by water utility managers. Their critique is that few people make a consciously economic decision when turning on the tap, and consequently they view the price elasticity of demand for water as being near zero. The D/C choice model directly allows for both economic and noneconomic influences: variation in behavior is due to both price and income (the fundamental economic determinants of behavior), and other forces represented by various sociodemographic variables, in a model which is utility theoretic.

[24]With a better handle on how other sociodemographic variables influence demand and perhaps better functional and distributional specifications, it may be possible to guarantee satiation with respect to water consumption.

However, it is costly to estimate the D/C choice model. A clear shortcoming of the analysis here is the sociodemographic specification, but better (that is, more) sociodemographic data is not a panacea, for more data implies more parameters to be estimated which in turn implies an exponential increase in MLE convergence time. Furthermore, experimentation with various sociodemographic forms implies estimating the model various times, again a costly solution. This suggests that some means of advance testing of model specifications would be useful. Note that model specification in the maximum likelihood framework includes not only distributional assumptions, but also functional form assumptions, both of the underlying indirect utility function and of the functional dependence on sociodemographic variables. These assumptions must be made in order to specify a likelihood model.

Finally, as has been alluded several times, there are significant costs associated with the D/C choice approach taken in this paper in terms of specificational assumptions which must be made to construct a two-error discrete/continuous choice model. Alternative approaches that are not fraught with such assumptions may prove to be quite fruitful. Semiparametric and nonparametric methods are a way to avoid such assumptions, and may show promise.

Figure 1

Utility Maximization in the D/C Choice Framework

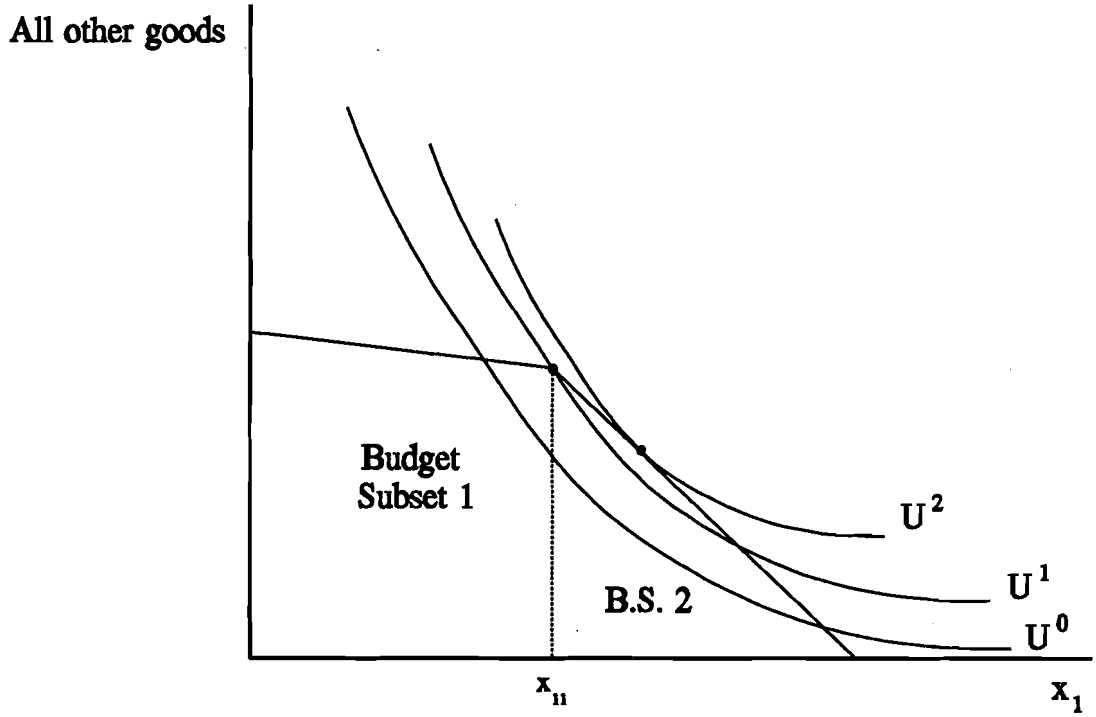
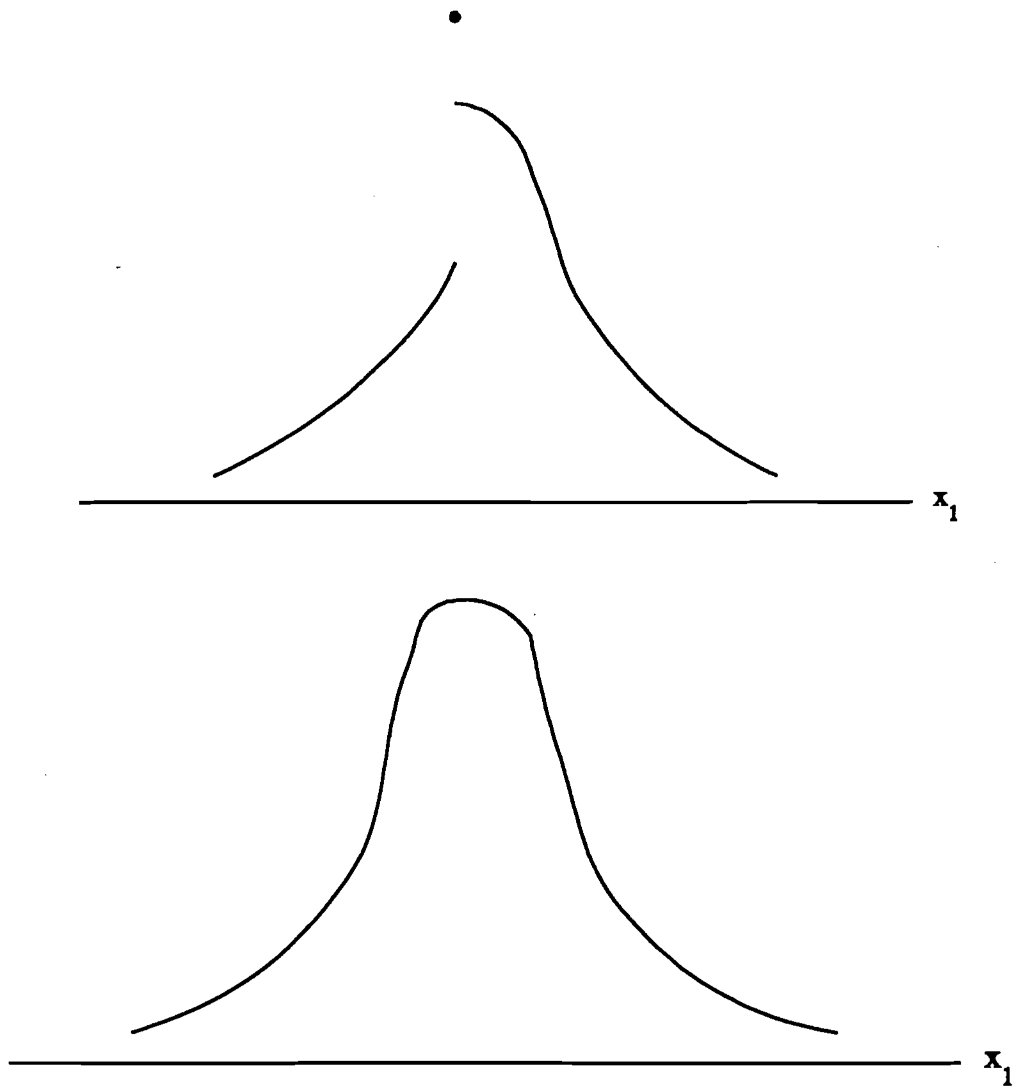


Figure 2.

Distributions of x_1:

(a) household heterogeneity only;* (b) both heterogeneity and perception error



* This figure shows a jump discontinuity which takes the form of an upwards spike in the frequency distribution of x_1. Depending on curvature properties of indifference maps, the rate structure and the distribution of ε, the jump discontinuity could occur either as an upward or downward spike.

Table 1. Descriptive Statistics for June to August Dataset

Variable Name	Units[a]	Means	Standard Deviations	Minimum	Maximum
x_1	1000 gals	14.966	12.435	0.1	212.8
y	$1000	1.156	0.485	0.137	2.967
lawn2	1000 sq ft	9.887	3.383	4.61	25.96
weather	inches	0.464	0.149	-0.1004	0.6324
bathrooms	number	1.635	0.520	1	3
house size	100 sq ft	18.337	5.288	4.44	36.11
days	number	30.508	2.786	4	62
p_{11}	$/1000 gals	1.206	0.090	1.09	1.34
p_{12}	$/1000 gals	1.442	0.147	1.25	1.67
x_{11}	1000 gals	20.000	0.000	20	20
d_1	$1000	-0.03346	0.00740	-0.04611	-0.00667
d_2	$1000	-0.02873	0.00630	-0.03963	-0.00019

[a] The variables x_1, y, weather, and days are all on a billing period rather than monthly or 30-day basis.

Table 2. Estimates of Water Demand Models

Variable	Model: OLS	Perception One-Error	Heterogeneity One-Error	Two-Error
constant	0.2989 (1.003)[a]	0.1807 (0.536)	0.2988 (0.92)	0.5184 (1.427)
lawn2	0.0105 (1.503)	0.0150 (2.087)	0.0105 (1.36)	0.0231 (2.728)
weather	0.2196 (1.544)	0.4331 (3.004)	0.2196 (1.18)	1.5945 (9.244)
bathrooms	0.1235 (2.289)	0.1154 (2.046)	0.1235 (2.39)	0.1336 (2.000)
house size	-0.0023 (-0.353)	0.0007 (0.104)	-0.0023 (-0.29)	0.0046 (0.598)
days	0.0334 (4.383)	0.0346 (4.131)	0.0334 (4.41)	0.0340 (3.636)
p_1	3.1963 (16.975)	2.5391 (13.070)	3.1962 (12.20)	-1.8989 (-6.421)
y + d	0.0976 (1.225)	0.1215 (1.422)	0.0976 (1.03)	0.1782 (1.864)
$\sigma_\varepsilon + \sigma_\eta$	0.809	—	—	—
σ_ε	—	—	0.8077 (104.37)	0.8737 (31.210)
σ_η	—	0.8102 (56.447)	—	0.3763 (8.229)
$\bar{R}^2$	0.209	—	—	—
F(7,1695)	65.419	—	—	—
Mean LL	—	-0.31589	-1.23177	-1.26417

[a] The numbers in parentheses are t-statistics, or asymptotic t-statistics in the maximum likelihood D/C choice models.

Table 3. Predicted Values of Consumption

Two-Error Model

Evaluated at:	Sample	$E_\varepsilon(x_1)$ Expectation of Planned Consumption	$E_{\varepsilon\eta}(x_1)$ Expectation of Observed Consumption	$\bar{x}_1(p_{11}, y+d_1; Z)$
Sample Mean	14.9663	14.9233	18.0443	12.0922
Sample Median	11.5000	14.9247	17.7379	11.8023

Table 4. Price and Income Elasticities

Two-Error Model

Evaluated at:	Price Elasticities		Income Elasticities	
	Expectation of Planned Consumption	Expectation of Observed Consumption	Expectation of Planned Consumption	Expectation of Observed Consumption
Sample Means	-1.586	-1.133	0.1543	0.1102
Sample Medians	-1.629	-1.233	0.1582	0.1197

APPENDIX

The log-likelihood functions for all three stochastic specifications are presented, though not derived below. For more detail on the derivation, see Hewitt (1993).

The perception-error-only likelihood function of (6) requires only the distributional assumption for η and the functional form of $\bar{x}_1$ to be complete. In log-likelihood form, it is:

$$\mathrm{LnL}_p = \sum_{\text{all}} \ln\left\{\frac{\exp\left(-w_1^2/2\right)}{\sigma_\eta}\left[1-\Phi(t)\right] + \frac{\exp\left(-w_2^2/2\right)}{\sigma_\eta}\Phi(t)\right\} \tag{A1}$$

where

$$\begin{aligned} w_k &= \left[\ln x_1 - Z\delta - \mu\ln(y+d_k) - \alpha_1 \ln p_{1k}\right]/\sigma_\eta \\ t &= \left[\ln x_1 - \ln x_{11}\right]/\sigma_\eta \end{aligned} \tag{A2}$$

and where Φ is the standard normal cdf.

Similarly, the heterogeneity-error-only likelihood function of (8) requires only the distributional assumption for ε and the functional form of $\bar{x}_1$ to be complete. In log-likelihood form, it is:

$$\mathrm{LnL}_h = \sum_{\text{block 1}}\left[-w_1^2/2 - \ln(\sigma_\varepsilon)\right] + \sum_{\text{kink}} \ln\left[\Phi(t_2) - \Phi(t_1)\right] + \sum_{\text{block 2}}\left[-w_2^2/2 - \ln(\sigma_\varepsilon)\right] \tag{A3}$$

where

$$\begin{aligned} w_k &= \left[\ln x_1 - Z\delta - \mu\ln(y+d_k) - \alpha_1 \ln p_{1k}\right]/\sigma_\varepsilon \\ t_k &= \left[\ln x_{11} - Z\delta - \mu\ln(y+d_k) - \alpha_1 \ln p_{1k}\right]/\sigma_\varepsilon \end{aligned} \tag{A4}$$

Note, however, that deriving a likelihood function from expression (9) or (13) requires knowledge of $f_{\varepsilon\eta}(\varepsilon, \eta)$ and $g_{\varepsilon+\eta,\varepsilon}(\varepsilon+\eta,\varepsilon)$. If ε and η are independent, then $f_{\varepsilon\eta}(\varepsilon, \eta) = f_\varepsilon(\varepsilon)$ $f_\eta(\eta)$. The middle expression of the rhs of (13) involves a single value of η but a range of ε values, or is of the form $f_\eta(\eta)F_\varepsilon(\varepsilon)$. The first and third expressions of the rhs of (13) involve a range of values of ε, and by construction then, a range of values of η though each value that ε takes on determines a single value of η Equation (10) is instrumental in the likelihood form of equation (13), but a greatly simplified expression of the likelihood results if we use the definitional relationship: $g_{v,\varepsilon}(v, \varepsilon) = g_{v|\varepsilon}(v \mid \varepsilon) f_{\varepsilon}(\varepsilon)$, where $v = \varepsilon + \eta$. When $g_{v,\varepsilon}$ is a jointly dependent normal distribution, its value over a range of ε values can be factored into standard normal pdf and cdf terms with appropriate transformations of variables. Letting ρ denote the correlation coefficient between v and ε, the log-likelihood for (13) is:

$$LnL_{h+p} = \sum_{all} \ln\left\{\frac{\exp\left(-w_1^2/2\right)}{\sigma_v}\Phi(r_1) + \frac{\exp\left(-v^2/2\right)}{\sigma_\eta}\left[\Phi(t_2) - \Phi(t_1)\right] + \frac{\exp\left(-w_2^2/2\right)}{\sigma_v}\left[1 - \Phi(r_2)\right]\right\} \qquad \text{(A5)}$$

where

$$\begin{aligned} w_k &= \left[\ln x_1 - Z\delta - \mu \ln(y + d_k) - \alpha_1 \ln p_{1k}\right]/\sigma_v \\ t_k &= \left[\ln x_{11} - Z\delta - \mu \ln(y + d_k) - \alpha_1 \ln p_{1k}\right]/\sigma_\varepsilon \\ r_k &= (t_k - \rho\, w_k)/\sqrt{1-\rho^2} \\ v &= \left[\ln x_1 - \ln x_{11}\right]/\sigma_\eta \\ \sigma_v &= \sqrt{\sigma_\varepsilon^2 + \sigma_\eta^2} \qquad \rho = \sigma_\varepsilon/\sigma_v \end{aligned} \qquad \text{(A6)}$$

Note that the $(2\Pi)^{-1/2}$ terms may be dropped without affecting the likelihood maximization.

REFERENCES

Burtless, Gary, and Jerry A. Hausman. 1978. "The Effect of Taxation on Labor Supply: Evaluating the Gary Income Maintenance Experiment." *Journal of Political Economy* 86(6):1101-1130.

Camp, Robert C. 1978. "The Inelastic Demand for Residential Water: New Findings." *Journal of the American Water Works Association* 70(8):453-458.

Chicoine, David L. and Ganapathi Ramamurthy. 1986. "Evidence on the Specification of Price in the Study of Domestic Water Demand." *Land Economics* 62(1): 26-32.

Chicoine, David L., Steven C. Deller, and Ganapathi Ramamurthy. 1986. "Water Demand Estimation Under Block Rate Pricing: A Simultaneous Equation Approach." *Water Resources Research* 22(6):859-863.

Deller, Steven C., David L. Chicoine, and Ganapathi Ramamurthy. 1986. "Instrumental Variables Approach to Rural Water Service Demand." *Southern Economic Journal* 53(2):333-346.

Foster, Henry S. Jr., and Bruce R. Beattie. 1979. "Urban Residential Demand for Water in the United States." *Land Economics* 55(1):43-58.

Hall, Robert H. 1973. "Wages, Income, and Hours of Work in the U.S. Labor Force." *Income Maintenance and Labor Supply*. Institute for Research on Poverty Monograph Series. pp. 102-162.

Hanemann, W. Michael. 1984. "Discrete/Continuous Models of Consumer Demand." *Econometrica* 52(3):541-561.

Hausman, Jerry A. 1981. "Labor Supply," in *How Taxes Affect Economic Behavior*, eds., Henry Aaron and Joseph Pechman, Washington, DC: Brookings Institution.

Hausman, Jerry A. 1985. "The Econometrics of Non-linear Budget Sets." *Econometrica* 53(6): 1255-1282.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153-161.

Hewitt, Julie A. 1993. *Watering Households: The Two-Error Discrete-Continuous Choice Model of Residential Water Demand.* Ph.D. Dissertation, University of California at Berkeley.

Hewitt, Julie A. and W. Michael Hanemann. 1994. "A Discrete/Continuous Choice Approach to Residential Water Demand Under Block Rate Pricing." Montana State University Department of Agricultural Economics and Economics Staff Paper No. 94-4.

Howe, Charles W., and F. P. Linaweaver, Jr. 1967. "The Impact of Price on Residential Water Demand and Its Relation to System Design and Price Structure." *Water Resources Research* 3(1):13-31.

Jones, C. Vaughan and John R. Morris. 1984. "Instrumental Price Estimates and Residential Water Demand." *Water Resources Research* 20(2): 197-202.

Lewbel, Arthur. 1985. "A Unified Approach to Incorporating Demographic or Other Effects into Demand Systems," *Review of Economic Studies* 52:1-18.

Martin, William E., Helen M. Ingram, Nancy K. Laney, and Adrian H. Griffin. 1984. *Saving Water In A Desert City*. Washington, D.C.: Resources for the Future.

Moffitt, Robert. 1986. "The Econometrics of Piecewise-Linear Budget Constraints: A Survey and Exposition of the Maximum Likelihood Method." *Journal of Business and Economics Statistics* 4(3):317-328.

Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes. 1974. *Introduction to the Theory of Statistics*, 3rd ed. New York: McGraw-Hill.

Neary, J. P., and K. W. S. Roberts. 1980. "The Theory of Household Behavior Under Rationing." *European Economic Review* 13: 25-42.

Nieswiadomy, Michael L., and David J. Molina. 1988. "Urban Water Demand Estimates Under Increasing Block Rates." *Growth and Change* 19(1):1-12.

Nieswiadomy, Michael L., and David J. Molina. 1989. "Comparing Residential Water Demand Estimates under Decreasing and Increasing Block Rates Using Household Data." *Land Economics* 65(3):280-289.

Nordin, John A. 1976. "A Proposed Modification of Taylor's Demand Analysis: Comment." *The Bell Journal of Economics* 7:719-721.

Pollak, Robert A. and Terence J. Wales. 1981. "Demographic Variables in Demand Analysis," *Econometrica*, 49 (6):1533-1551.

Pudney, Stephen. 1989. "Modelling Individual Choice: The Econometrics of Corners, Kinks, and Holes." Basil Blackwell, Oxford.

Reece, W. and K. Zieschang. 1985. "Consistent Estimation of the Impact of Tax Deductibility on the Level of Charitable Contributions." *Econometrica* 53(2):271-294.

Rothbarth, E. 1941. "The Measurement of Changes in Real Income under Conditions of Rationing." *Review of Economics Studies* 8(2):100-107.

Taylor, Lester D. 1975. "The Demand for Electricity: A Survey." *The Bell Journal of Economics* 6(1):74-110.

U.S. Department of Commerce. 1989. *Business Statistics: 1961-1988*, 26th edition, A Supplement to the Survey of Current Business, Bureau of Economic Analysis (December).

ESTIMATION OF WILLINGNESS TO ACCEPT COMPENSATION FOR PUBLIC AND PRIVATE GOODS FROM THE CHOOSER REFERENCE POINT BY THE METHOD OF PAIRED COMPARISON

George L. Peterson, Thomas C. Brown, Daniel W. McCollum, Paul A. Bell, Andrej A. Berjulin, and Andrea Clarke*

***The authors' affiliations are, respectively the Rocky Mountain Forest and Range Experiment Station, USDA Forest Service, Fort Collins, CO (first three authors), and Department of Psychology, Colorado State University, Fort Collins, CO (second three authors).**

ESTIMATION OF WILLINGNESS TO ACCEPT COMPENSATION FOR PUBLIC AND PRIVATE GOODS FROM THE CHOOSER REFERENCE POINT BY THE METHOD OF PAIRED COMPARISON

INTRODUCTION

The subject of this paper is hypothetical contingent market estimation of willingness to accept compensation (WTA) for nonmarket goods and services. The purposes are (1) to pilot test a psychometric method of contingent estimation of WTA, and (2) to use that method to test several hypotheses about human valuation behavior. These hypotheses include (a) that people are willing and able to make choices among gains in nonmarket goods and gains in money or market goods, (b) that their choices obey the postulate of rationality, and (c) that the implied equivalent variation values are plausible and conservative.

Economists to date have been unable to obtain a valid conservative estimate of WTA for the loss of a nonmarket good using the contingent valuation (CV) method. Although WTA may be the appropriate measure of monetary value in some cases, e.g., for ex post assessment of damages from environmental accidents, theory and experience demonstrate that estimates of WTA obtained by CV are sensitive to framing and context effects and likely to exhibit significant biases. Researchers have therefore largely abandoned CV estimates of WTA and generally rely on estimates of willingness to pay to avoid the losses in question, although the use of CV to estimate WTP for nonmarket goods is also not without controversy (e.g., Cambridge Econ, Inc. 1992).

In response to the controversy surrounding CV, Arrow et al. (1993) call for methods that yield conservative estimates and include appropriate allowance for the range of available substitutes, among other things. In this paper we attempt to develop a hypothetical market method for estimating WTA that meets those requirements.

We attempt to satisfy the conservative requirement for WTA estimation by placing the respondent in the position of "chooser" rather than "seller" (Kahneman, Knetsch, and Thaler 1990). This approach avoids some of the experimental and behavioral phenomena (e.g., loss aversion) that otherwise might increase the difference between WTA and WTP beyond income effects[1]. We attempt to incorporate a range of available substitutes by requiring the respondent to make multiple discrete tradeoff choices among several nonmarket goods, sums of money, and private goods. We hypothesize that when a public good is embedded in a context requiring choices between it and the elements of a sufficiently diverse set of private goods and sums of money, the consumer will give adequate consideration to substitution and income (opportunity cost) effects as called for by Arrow et al (1993).

The method also incorporates a test of the transitivity axiom, that is, that people can exchange money and the goods and services that money can buy for nonmarket goods and services on a unidimensional monetary continuum. This axiom lies at the root of any monetary valuation of public goods, whether WTP or WTA. Finally, if the method is successful, it may provide a framework from which to develop standard methods for estimating WTA.

The first section following this introduction presents a brief summary of the well-known economic theory. The second explains the measurement method in terms of psychological choice theory and psychometric methods, and integrates the economic and psychological concepts into a unified approach to monetary valuation. The third section describes the experimental design, including the hypotheses to be tested and the experimental methods used. The fourth section reports the results of pilot studies of several public and private goods. The final section of the paper summarizes and discusses our findings, including recommendations for future research.

The research reported here is exploratory. The results are intended only to be illustrative of the kinds of data, analyses, and conclusions enabled by the method. Specific numerical results are tentative at this time and should not be generalized.

---

[1] Some authors argue that WTA is, by definition, a behavioral phenomenon, not a theoretical construct. If loss aversion is real, for example, it is by argument of consumer sovereignty a legitimate part of WTA (Kahneman et al. 1990). Economic theory, they say, is based on a premise of global rationality that is not descriptive of the bounded rationality of human choice behavior (Simon 1985). This question remains controversial among many economists, however, who argue that we must define WTA by economic theory, that observed differences between WTA and WTP beyond theoretical limits are artifacts of poor measurement, not valid assigned values, or that there is no property right to the added value. We make no attempt here to resolve these controversies.

ECONOMIC THEORY

We are concerned with two aspects of economic theory: (1) the applicability of neoclassical microeconomic consumer theory to choices involving public goods, and (2) the theory of welfare change, on the assumption that the consumer theory is applicable.

Utility Theory and the Nature of Public Goods

Neoclassical microeconomic consumer theory states that a consumer's preferences can be described by an ordinal utility function. Necessary assumptions include (1) the consumer knows whether he prefers A to B, B to A, or is indifferent between the two for all possible pairs of alternatives; (2) only one of the three possibilities is true for any pair of alternatives; and (3) if he prefers A to B and B to C, he will prefer A to C. These three conditions constitute the postulate of rationality, which requires the consumer be able to rank commodities in order of preference. It is not necessary to assume that the consumer holds a cardinal measure of utility. The much weaker assumption of consistent ranking of preferences is sufficient (Henderson and Quandt 1980). These assumptions define the boundaries of neoclassical microeconomic consumer theory and reduce the consumer's choice problem to the budget constrained maximization of utility, the foundation on which CV lies (Deaton and Muellbauer 1981).[2]

One of the criticisms of CV identified by Arrow et al. is "inconsistency with rational choice." If consumers' choices among public goods and private goods or among public goods and sums of money violate the axioms of utility theory, those choices are not consistent with economic theory, and monetary valuation derived therefrom is not valid as far as that theory is concerned. We ask (1) are consumers willing to make such choices, (2) if willing, are they able, and (3) if willing and able, do the choices produce a consistent ranking of alternatives?

[2]Deaton and Muellbauer (1981) define the required axioms of utility theory as reflexivity, completeness, transitivity, continuity, and nonsatiation.

Critics of CV argue from the perspective of the global rationality of neoclassical microeconomic consumer theory that if CV produces results that are inconsistent with rational choice, then CV must be at fault, because the theory states that choices must be rational. The fault may be more fundamental, namely, (1) that we are asking people to violate the axioms of utility theory by the kinds of choices we ask them to make, and/or (2) that people simply cannot or do not behave with global rationality. These problems, if they exist, are not unique to CV. Perhaps a more fundamental question is whether it is possible to identify a public good in an hypothetical contingent market context such that respondents know what it is and can think of it in terms of monetary exchange.

The question of transitivity requires some background. Hicks (1956, p 166) suggested " . . . we ought to think of the consumer as choosing, according to his preferences, between certain *objectives*; and then deciding, more or less as the entrepreneur decides, between alternative *means* of reaching those objectives." He thus saw demand for goods as *derived* from demand for the objectives served by those goods, with the goods being employed as input factors in a production process aimed at achieving the objectives.

Morishima (1959) translated Hicks' verbal theory into the mathematical language of Slutsky and used a linear programming approach instead of the traditional marginal theory of the firm to understand intrinsic complementarity between goods. These two authors laid a foundation for the household production theory of Becker (1965) and Lancaster's new approach to consumer theory (1966).

Following this logic, we assume that when confronted with a choice among several things, the consumer has several objectives in need of satisfaction. These objectives motivate demand for goods. Each alternative good possesses more or less of several properties that make it more or less effective as an input factor in production processes by which the consumer satisfies the objectives. The consumer's choice thus includes four components: (1) choice (or weighting) of objective(s), (2) choice of production process, (3) choice (or weighting) of characteristics, and (4) choice of alternative. We focus hereafter on the objectives, the goods, and their attributes.

We assume that when a consumer faces a choice between a public good and a sum of money (or private good), she must first establish the framework of objectives. The choice may be for a single objective, more than one compatible objective, or two or more contradictory objectives. If the consumer desires to satisfy several noncontradictory objectives, she might (but does not necessarily) combine the objectives into a unidimensional criterion by means of some weighting function[3].

Different goods will have different capacities to satisfy these objectives because of their different properties. Considering all these things, the consumer maximizes utility, subject to the budget constraint, by choosing among goods. In order for theoretically consistent monetary valuation to be possible, the utility function must be a consistent and transitive ranking of the alternatives. We note, however, that the utility function is the outcome, not the cause, of the choice. Depending on how the consumer goes about the decision process, that utility function may or may not have the required properties[4].

Economic theory clearly supports such a complex decision process. The question is whether people behave the way theory developed to date requires them to behave in order for their choices to be transitive. For example, if the consumer fixes on different objectives or different properties or changes the decision rules when comparing A with B, B with C, and A with C, the result could be an apparent inconsistent ordering of preferences.

Monetary Valuation of Public Goods

If we assume that monetary valuation of public goods does not violate the axioms of utility theory, we can invoke the standard economic theory as in Figure 1 (Freeman 1979). In Figure 1 the vertical axis measures the monetary value of the consumer's endowment of private goods, including money, and

[3]Peterson and Worrall (1970) show that contradictory objectives may produce satisficing, rather than utility maximizing, behavior.

[4]Economic theory and the utility function are merely attempts to describe the process by which an individual chooses among a set of alternatives. Our objective is not to debate how well economic theory describes that process, but to design an experiment in which individuals can be observed in making choices among public goods, private goods, and sums of money. The outcomes of those choices can then be tested for consistency with theory.

the horizontal axis measures the public good endowment. U_1 and U_2 are isoquants of an individual consumer's utility function describing the trade-off relationship between public and private goods.

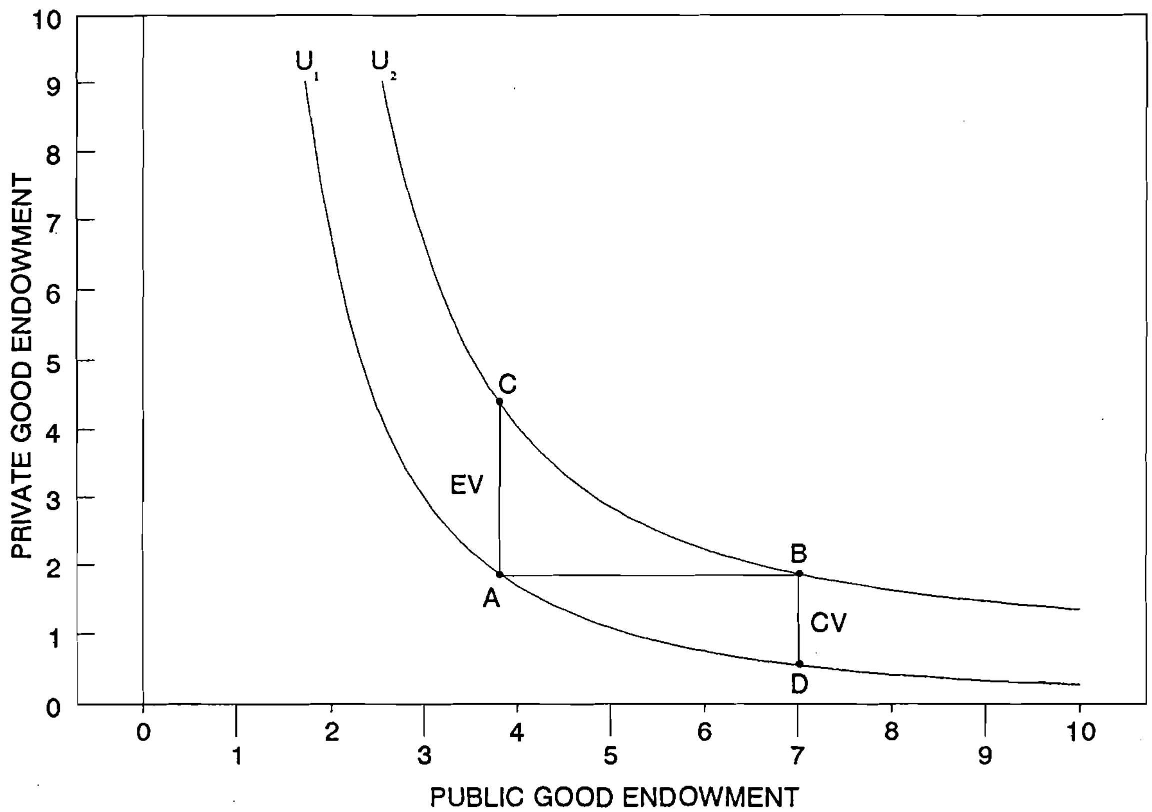
Assume the consumer is a "chooser" (Kahneman, Knetsch, and Thaler 1990) at reference point A on U_1 and faces a choice between (1) an increase in the endowment of public goods to B on U_2 and (2) an increase in private goods (and/or money) to C on U_2. Because B and C lie on the same isoquant, the consumer is indifferent between the alternatives. AC measures equivalent variation (minimum WTA) for the gain AB, while BD measures compensating variation (maximum WTP) to obtain the gain AB. WTA > WTP if there is a significant income effect of the change in public good endowment.

An important controversy surrounding WTA, however, is "loss aversion," the idea that the pain of loss is greater than the pleasure of gain, other things being equal. Empirical contingent valuation estimates of WTA generally exceed WTP by an amount greater than the income effect consistent with economic theory, and loss aversion is one of several plausible explanations (Gordon and Knetsch 1979; Tversky and Kahneman 1981; Knetsch 1984; Kahneman and Tversky 1984; Fisher et al. 1988; Kahneman, Knetsch, and Thaler 1991). Figure 2 moves the consumer to reference point B as a seller or loser, rather than a chooser. According to prospect theory and loss aversion, reduction of the public good endowment by an amount equal to BA requires the consumer to move to point E on U_3 in order to maintain indifference, rather than to point C on U_2. Under this theory, the correct measure of minimum WTA is AE, not AC, and the utility function has a discontinuity at B.[5]

Research has shown that WTA tends to be greater for sellers than for choosers (Kahneman, Knetsch, and Thaler 1990), and placing the consumer at reference point A as a chooser apparently avoids the question of loss aversion. Thus, a method that estimates WTA from the chooser reference point should give a more conservative result than an estimate from the seller or loser reference point.

Discovery of the indifference isoquants is a difficult if not impossible task. Assume, however, that we can define a set of incentive compatible choices that include several relevant private goods,

[5]It is "correct" in the sense that AE measures the true monetary value of the de facto felt loss under the loss aversion theory. Whether the consumer loses a de jure property right when the public good endowment changes from B to A is an issue beyond the scope of this paper.



U1
U2
C
EV
B
A
CV
D
PRIVATE GOOD ENDOWMENT
PUBLIC GOOD ENDOWMENT

FIGURE 1

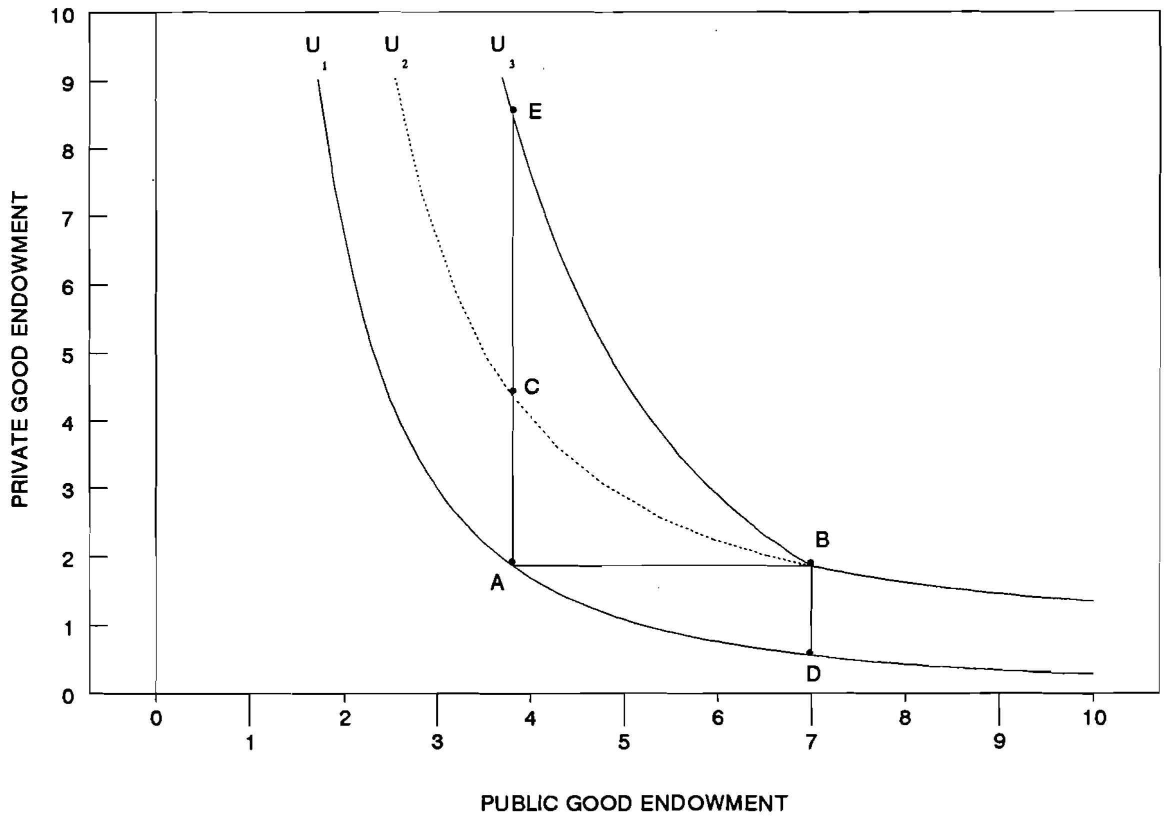

U1
U2
U3
E
C
B
A
D
PRIVATE GOOD ENDOWMENT
PUBLIC GOOD ENDOWMENT
0
1
2
3
4
5
6
7
8
9
10

FIGURE 2

several sums of money, and the public good(s) in question. It is then only necessary to order the consumer's preferences among the elements of this set in order to bound WTA for the public good(s). If we select the private goods and sums of money judiciously, we can "capture" the public good(s) within narrow boundaries, thereby obtaining a reasonably precise estimate of WTA.

PSYCHOLOGICAL THEORY

To develop an experiment by which to obtain an ordering of preferences among the elements of a set containing public goods, private goods, and sums of money, we turn to psychology and psychometric methods. Psychologists have, for generations, been developing and applying methods for ordering preferences. Too often academicians like to remain within the walls of their own disciplines, however, as lamented by Kenneth Boulding (1988):

> ". . . economics and psychology . . . are continents of the mind separated by a very wide ocean, no doubt produced by academic continental drift. Furthermore, they seem to be continents without any good harbors. . . . It is a fundamental principle of economics that specialization without trade is worthless. Unfortunately, in the continents of the mind, specialization seems to feed on itself, and there are large, invisible tariff barriers against the interchange of ideas."

Valuation is a psychological phenomenon, and we believe it is time to remove the tariff barriers, build some safe harbors, and establish a free trade agreement between economics and psychology. In this experiment we turn to the method of paired comparison as a way to obtain an ordinal ranking of public goods, private goods, and sums of money from the chooser reference point.

Why Use Paired Comparisons?

The method of paired comparisons (Fechner 1860; Thurstone 1927; Guilford 1954; Edwards 1957; Torgerson 1958; Bock and Jones 1968; David 1988; Kendall and Gibbons 1990) is a well-developed and established psychometric method for ordering preferences among the elements of a choice set. Given a set of t objects, the method presents them independently in pairs as $(t/2)(t-1)$ discrete binary choices[6]. The respondent simply chooses the preferred item in each pair. If there are no preference errors, and if the preferences obey the axioms of utility theory, the result will be a perfect rank ordering of the objects.

Why not simply ask the respondent to arrange the objects in rank order? The answer is that the method of paired comparisons allows intransitivity, whereas rank order does not. Preference intransitivity occurs in the form of circular triads, such as A>B>C>A. To quote Tversky (1969) as found in David (1988 pp 3-4),

> ". . . a circular triad denotes an inconsistency on the part of the judge, and its simplest explanation is that the judge is at least partially guessing when declaring preferences. The judge may be guessing because of incompetence or because the objects are in fact very similar . . . But guessing is not the only explanation, for there may be no valid ordering of the three objects even when they differ markedly. Their merit may depend on more than one characteristic, and it is then somewhat artificial to attempt an ordering on a linear scale. Under these circumstances the judge must mentally construct some function of the relevant characteristics and use this as a basis for comparison. It is not surprising that in complicated preference studies the function is vague and may change from one paired comparison to the next, especially when different pairs of objects may cause the judge to focus attention on different features of the objects. This last point helps to account for situations where a

[6]When the number of objects is large, various methods can be used to reduce the number of choices.

particular circular triad occurs frequently in repetitions of the experiment. However, circularity can occur even with a well-defined preference criterion based on two or more underlying dimensions (Tversky 1969)."

Application of the method of paired comparisons to economic choices is simply a special binary case of the CV method of discrete choice advocated by Arrow et al. (1993). However, application of the paired comparison method to economic choices in the manner we propose offers at least six advantages. First, it integrates into the economic valuation problem more than a hundred years of psychometric research on measurement of preferences for subjective stimuli, including rigorous probability theory, an arsenal of statistical tests, and applied experience dating back to Fechner (1860).

Second, paired comparisons can incorporate direct tradeoffs among public goods, market-priced private goods, and sums of money, thereby providing the respondent with strong incentives to consider income and substitution effects when making choices, thus improving the incentive-compatibility of the experiment. Third, it allows a test of the hypothesis that the individual's decision making behavior complies with the transitivity axiom of utility theory. Fourth, it provides individually reliable estimates of preference ordering (and of assigned value), thereby enabling in-depth analysis of individual behavior while enhancing the ability to identify market segments and calibrate the estimate of value to the extent of the appropriate market. Fifth, it allows convenient perturbation of the reference point, context, and frame of the experiment to further test hypotheses about economic choice behavior. Sixth, it offers promise for development of simple standard methods for estimating WTA.

As with any method, paired comparisons requires effective specification of the goods for which we require estimates of WTA, or in the words of Arrow et al. (1993), "Adequate information must be provided to respondents about the environmental program that is offered. It must be defined in a way that is relevant . . .". Herein lies a formidable challenge for any method, and the method of paired comparisons does not avoid this challenge.

As previously stated, the method of paired comparison allows a test of the hypothesis that an individual's preferences among a set of choices are transitive. If the individual's observed intransitivity can be attributed to error variance alone, rather than to systematic effects, and if the data comply with certain assumptions, repeated responses from a given individual or individual responses from a group of similar individuals yield a preference order among the objects in the choice set. Under certain conditions, it is also possible to derive an interval scale of preference "magnitude" that preserves the ordinal relationship. Kendall and Gibbons (1990) and David (1988) describe the applicable probability theory and statistical tests. Edwards (1957), Torgerson (1958), and Bock and Jones (1968) explain the analytical methods and underlying assumptions for interval scale estimation. Maxwell (1974) provides a simplifying analytical procedure based on the logistic transformation.

The psychometrics of paired comparison has a counterpart in economics in the form of utility maximizing discrete choice theory. Luce (1959, 1977) formalized Arrow's (1951) "independence of irrelevant alternatives" (IIA) assumption into a choice axiom. This model has been shown to be essentially equivalent to Thurstone's (1927) "law of comparative judgment" (Case V), if Thurstone's assumption of independent, normally distributed random variables is replaced by one of double exponential, random disturbances (Yellott 1977, McFadden 1973). The difference distribution of two independent double exponential random variables is the logistic distribution, which is the basis for the multinomial logit model. This intertwining among the roots of economic and psychological choice theories offers an opportunity to cast paired comparison in terms of utility maximizing discrete choice theory, thus providing an economically consistent justification for application of the well-developed psychometrics of paired comparison to CV.

EXPERIMENTAL DESIGN

The experiment reported here builds on numerous pilot studies that helped develop and refine concepts and methods used in the experiment reported here. In this experiment we use 21 stimuli, consisting of six locally relevant public goods, four private goods, and eleven sums of money. Because paired comparison of two unequal amounts of money is trivial, we take that section of the matrix as given and present respondents only with choices between public goods, between private goods, between public and private goods, between public goods and sums of money, and between private goods and sums of money. Each respondent thus makes choices for 155 pairs. Two hundred twenty-one respondents participated in the study for a total of 32,550 binary choices. The respondents were students at Colorado State University.

Table 1 lists the goods and gives brief descriptions.[7] The eleven sums of money were $1, $25 to $100 in intervals of $25, and $100 to $700 in intervals of one hundred dollars. The public and private goods used in the experiment were derived from earlier pilot studies in order to get good variation and distribution across the dollar magnitudes.

The experiment was administered by means of a computer code that presented the stimuli on the screen in random order for each respondent. We gave the public and private goods short names, and these short names appeared side-by-side on the screen, with their position (right versus left) also randomized. The respondent entered a choice by pressing the right or left arrow key and could correct mistakes by pressing "backspace." At the end of the 155 paired comparisons, the computer code repeated in random order those pairs that were not consistent with the dominant rank order. It also randomly selected ten pairs that were consistent and repeated them. The computer then presented each respondent with a set of attitudinal and informational debriefing questions in a quantitative response format.

The computer program recorded (1) the choice for each pair in an ordered matrix, (2) the time in seconds required for each choice, (3) the sequence of each choice, (4) the pairs that were inconsistent with the dominant rank order, (5) all circular triads, (6) the number of times each good or sum of

[7]The descriptions given in Table 1 are very brief, which means that respondents probably do not have standardized or detailed perceptions of what they represent. Rigorous application of the method would require well developed and tested information scenarios. Whether it is possible to create adequate descriptions of public goods in a hypothetical market context is an important and unanswered question.

TABLE 1

PRIVATE GOODS

1. A meal at a Fort Collins restaurant of your choice, not to exceed $15. **(Meal)**

2. Two tickets and transportation to one of the following:
   A) A Colorado ski area of your choice.
   B) A concert of your choice in Denver (Contemporary or Classical).
   C) A Broncos, Rockies, or Nuggets game.
   D) A cultural event of your choice at the Denver Center for the Performing Arts.

   Estimated value: $75 **(Tickets)**

3. A nontransferable $200 certificate for clothing at a Fort Collins store of your choice. **(Clothes)**

4. A nontransferable certificate for you to make $500 worth of flights on an airline of your choice. **(Air Travel)**

PUBLIC GOODS

1. A no-fee library service that provides videotapes of all course lectures so that students can watch tapes of lectures for classes they are not able to attend. **(Videotape Service)**

2. Parking garages to increase parking capacity on campus such that students are able to find a parking place at any time, without waiting, within a five-minute walk of any building at no increase in the existing parking permit fee. **(Parking Capacity)**

3. Purchase by CSU of 2,000 acres of land in the mountains west of Fort Collins as a wildlife refuge for animals native to Colorado. **(Wildlife Refuge)**

4. A CSU-sponsored, on-campus springtime weekend festival with a variety of live music and student participation events with no admission fee. **(Spring Festival)**

5. Expansion of the eating area in the Lory Student Center to ensure that any student can find a seat at any time. **(Eating Area)**

6. A cooperative arrangement between CSU, local business groups, and the citizens of the community that would ensure the air and water of Fort Collins would be at least as clean as the cleanest 1% of the communities in the U.S. **(Clean Arrangement)**

TABLE 2
THREE EXPERIMENTAL SCENARIOS

SCENARIO	RESPONSIBILITY FOR OUTCOME	BENEFIT OF PVT. GOODS AND MONEY	N
1.	SHARED	SHARED	72
2.	SOLE	SHARED	76
3.	SOLE	SOLE	73

money occurred in circular triads, (7) preference switches for the inconsistent pairs and for the ten randomly sampled consistent pairs, and (8) responses to the debriefing questions.

Respondents had Table 1 in front of them at all times during the experiment and were free to refer to it at any time. Average total time to complete the 155 paired comparisons was about 10 minutes, not including the time required to become familiar with the instructions.

In order to test context effects, we presented three different choice scenarios to three independent subgroups of the 221 respondents. Table 2 lists the three scenarios and their respective sample sizes. The first scenario (SHARE-SHARE) asked the respondent to cast votes as if in a referendum, thus sharing responsibility for the outcome with all other students. This scenario specified that if a sum of money or private good "won" the election, all students would receive it. In the second scenario (SOLE-SHARE) the respondent had sole responsibility for the outcome, but if the outcome was a private good or sum of money, all students would share the benefit. The third scenario (SOLE-SOLE) gave sole responsibility for the outcome and specified that the respondent would be the sole recipient of the benefit if the choice was a private good or sum of money.

RESULTS

The results presented here are intended only to illustrate the kinds of analyses and findings enabled by the paired comparison method. They must be considered preliminary and must not be generalized beyond the sample. More detailed and rigorous analyses will be forthcoming in subsequent publications.

Transitivity

Figure 3 shows the degree of transitivity achieved by the 221 respondents. The maximum possible number of intransitive triads in this experiment is 330. We calculate a coefficient of consistency by subtracting the individual respondent's number of circular triads from 330, dividing the result by 330, and then multiplying by 100. Figure 3 shows that 90% of the respondents were at least

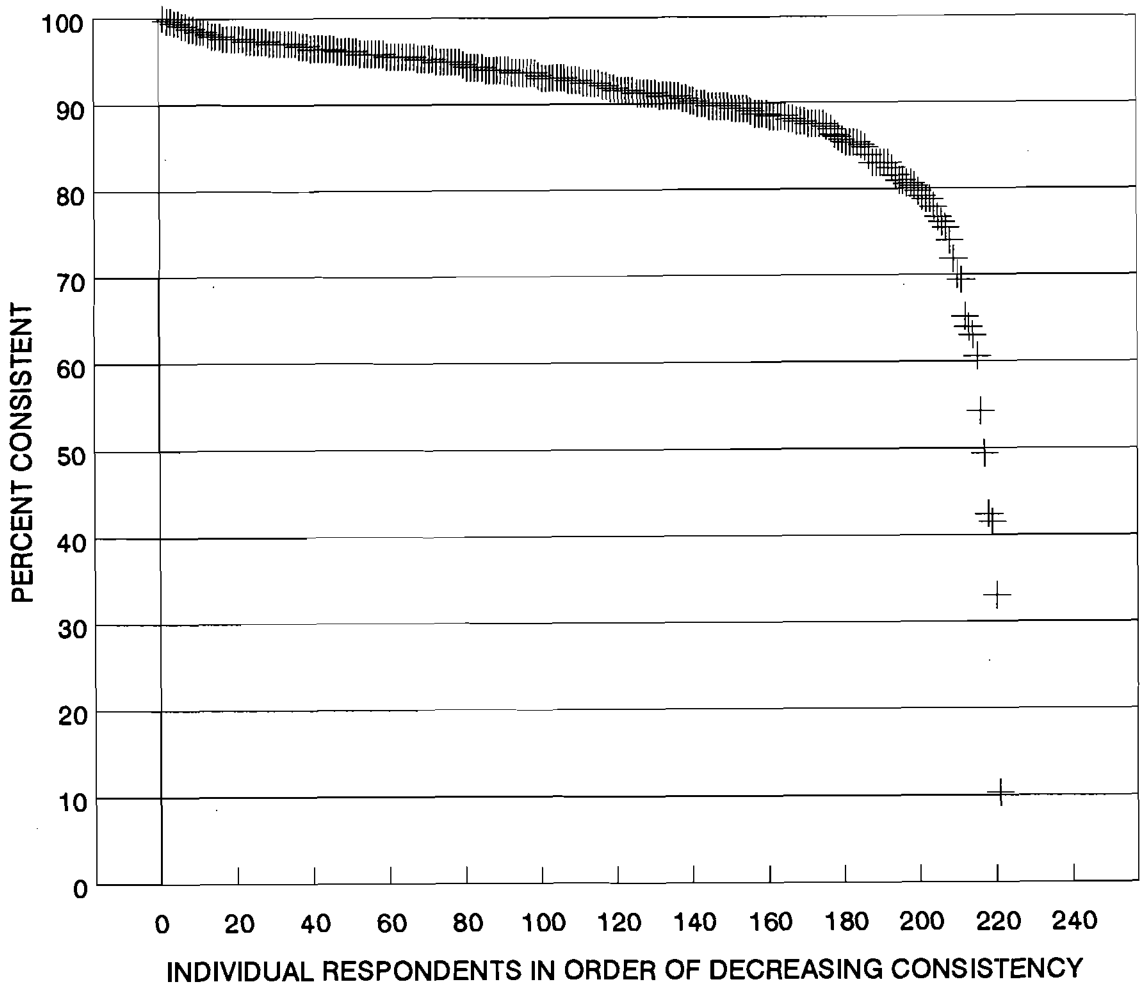


FIGURE 3
CONSISTENCY OF RESPONSE

80% consistent, and 68% were at least 90% consistent. We must now ask whether the observed intransitivity is random or systematic. This result suggests that most respondents are highly consistent, and that the observed inconsistency does not exceed what might be caused by random errors or indifference in some choices.

Figure 4 compares preference switching behavior for consistent and inconsistent pairs. An inconsistent pair is one for which the expressed preference is not consistent with the dominant rank order, thereby being a source of inconsistent triads. Recall that the computer code identifies all inconsistent pairs and repeats them in random order at the end. The program also repeats ten randomly selected consistent pairs. If random error and indifference are the causes of the inconsistency, we would expect the probability of switching for inconsistent pairs to be 0.5. Because the repeated consistent pairs are sampled from the whole matrix with the inconsistent pairs removed, we would expect the probability of switching to be less than 0.5.

The two distributions in Figure 4 are significantly different no matter what statistical test one chooses to use. For example, a simple student's t for the difference between the two means has a magnitude of 23.4 with 440 degrees of freedom. Clearly, respondents switch preference for inconsistent pairs more frequently than for consistent pairs. The cause of inconsistency may be mistakes which the respondent tends to correct under repetition, similarities that are too close to call consistently, or revision of decision rules during the course of the experiment. It is worth noting, however, that the average proportion of inconsistent pairs switched is close to 0.5 whereas the average proportion of consistent pairs switched is much less.

Preference Variance Within Goods

Another important question is whether the goods and sums of money all have the same variance across subjects. Each respondent's preferences yield a dominance score for each good, measured as the number of other goods dominated by the good in question. A highly preferred good will have a high dominance score, and vice versa. If all respondents order the goods identically and error variance is

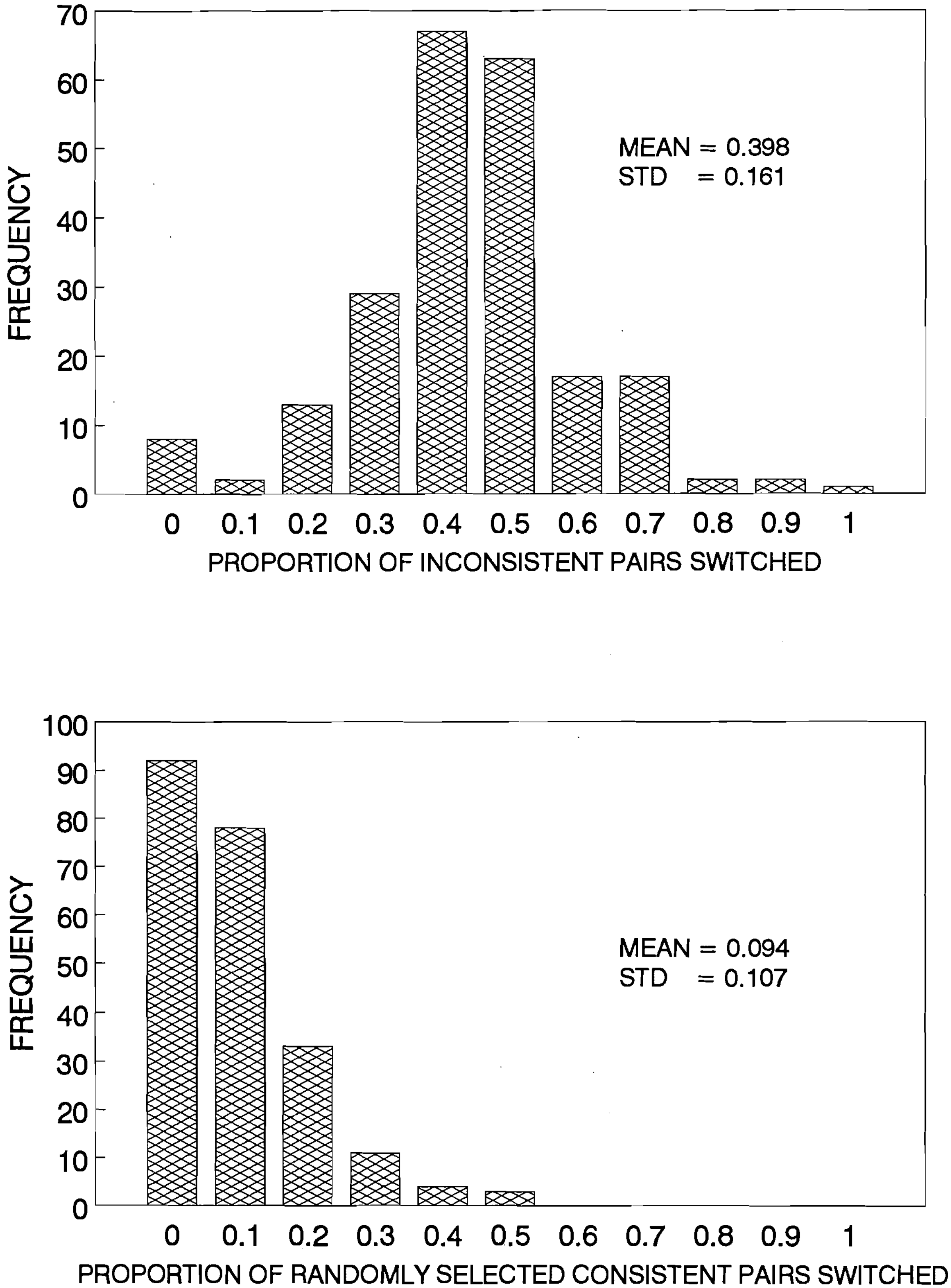


FIGURE 4
PREFERENCE SWITCHING BEHAVIOR

homogeneous, the goods will all have the same variance across subjects. If subjects disagree more about some goods than about others, or if they have greater error variance for some than for others, different goods will have different dominance profiles and different variance across subjects.

Figure 5 shows graphically the dominance profile variance for each good and sum of money across all 221 respondents, in descending order. Note that the public goods tend to have greater variance than the private goods. Wildlife Refuge, Clean Arrangement, Videotape Service, and Parking Capacity show the greatest variance, while Tickets, Clothes, and Meal show the least variance.

This variance pattern could be caused by greater disagreement among respondents about the public goods or greater error variance within respondents for the public goods. Greater error variance for public goods might be caused by poor or vague definition, greater difficulty in making choices that involve public goods, or poorly defined preferences. The data raise interesting questions about the observed phenomenon, but do not reveal the cause.

Estimation of Monetary Magnitudes

Monetary magnitudes for the public and private goods can be estimated in one of several ways: (1) bracketing by preference order, (2) linear interpolation within the brackets by dominance score, (3) derivation of an anchored interval scale by the psychometric law of comparative judgment, and (4) derivation of a mathematical function from the relationship between sums of money and either dominance scores or scale magnitudes.

Table 3 illustrates application of the first two methods, partitioned by the three context scenarios. The "mean score" in Table 3 is the mean dominance score, i.e., the average number of times the good or sum of money in question dominates other goods or sums of money. The aggregate preference order defined by the dominance scores brackets the public and private goods between sums of money also used in the experiment. The dominance scores have then been used to obtain finer estimation by linear interpolation within the brackets. It is useful to note that the paired comparison data allow preference ordering and monetary estimation by bracketing and interpolation for both individual and aggregate data.

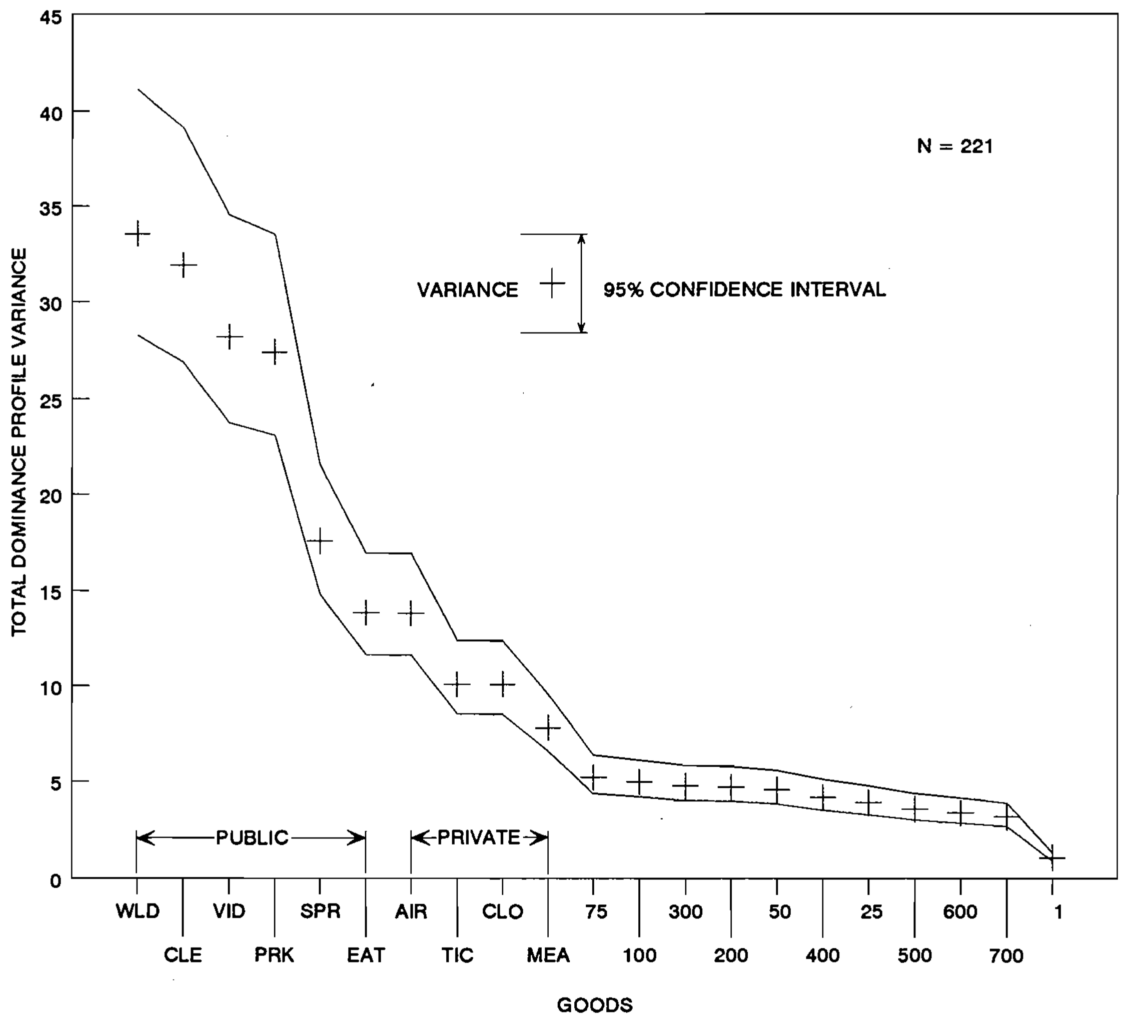


FIGURE 5
PREFERENCE VARIANCE WITHIN GOODS

TABLE 3 MEAN DOMINANCE SCORES AND DOLLAR ESTIMATES BY TREATMENT								
TREATMENT								
SHARE-SHARE			SOLE-SHARE			SOLE-SOLE		
GOODS	MEAN SCORE	$	GOODS	MEAN SCORE	$	GOODS	MEAN SCORE	$
$700	18.79	$700	$700	18.13	$700	$700	18.49	$700
$600	17.82	$600	$600	17.00	$600	$600	17.44	$600
$500	16.53	$500	$500	15.74	$500	$500	16.16	$500
$400	15.17	$400	WLD	14.41	$420	$400	14.79	$400
AIR	13.60	$301	$400	14.08	$400	AIR	14.38	$375
$300	13.58	$300	AIR	13.92	$386	WLD	14.34	$373
$200	12.06	$200	$300	12.91	$300	$300	13.15	$300
WLD	11.83	$191	CLE	12.88	$298	CLE	11.97	$220
CLO	11.10	$164	$200	11.12	$200	$200	11.68	$200
CLE	10.25	$131	CLO	10.50	$172	CLO	10.77	$167
VID	9.61	$107	VID	9.43	$124	TIC	9.12	$107
$100	9.42	$100	PRK	9.38	$122	$100	8.93	$100
TIC	9.26	$ 98	$100	8.89	$100	VID	8.19	$ 88
$ 75	7.60	$ 75	TIC	8.64	$ 97	SPR	7.70	$ 81
PRK	7.44	$ 73	SPR	7.82	$ 87	$ 75	7.33	$ 75
SPR	6.33	$ 56	$ 75	6.87	$ 75	PRK	7.26	$ 74
$ 50	5.89	$ 50	$ 50	5.28	$ 50	$ 50	5.36	$ 50
EAT	4.85	$ 36	EAT	5.21	$ 49	EAT	5.01	$ 45
MEA	4.35	$ 30	MEA	3.82	$ 29	MEA	3.85	$ 28
$ 25	3.99	$ 25	$ 25	3.54	$ 25	$ 25	3.66	$ 25
$ 1	0.54	$ 1	$ 1	0.43	$ 1	$ 1	0.40	$ 1

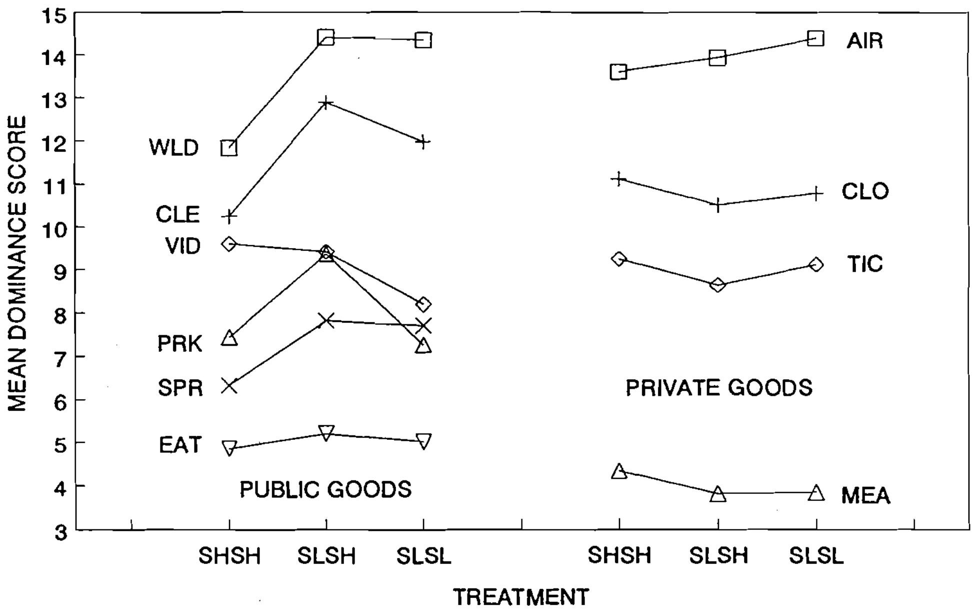
Application of psychometric scaling awaits further analysis to evaluate compliance of the data with the required underlying assumptions. We suspect, however, that such scaling will not change the results significantly.

Differences Among the Three Treatments

Table 3 shows that average preference order differs somewhat among the three treatments, and Figure 6 shows the mean dominance score profile for each good and sum of money across the three treatments. On face value, differences appear to be of greater magnitude and different pattern for public goods than for private goods and sums of money. With the exception of Video Service, the public goods show a consistent pattern: sole responsibility for shared private benefit drives the score higher than shared responsibility for shared private benefit. And, for all public goods, sole responsibility for sole private benefit drives the value lower than sole responsibility for shared private benefit. With the exception of Air Travel and $1, these effects are opposite for the private goods and sums of money. The binomial sign test shows these profile patterns to be significant. Analysis of variance also shows that mean dominance scores differ significantly across the three treatments for Parking Capacity, Wildlife Refuge, Spring Festival, Clean Arrangement, $200, $400, $500, $600, and $700. These findings encourage the speculative hypothesis that differences in moral responsibility may be at work and/or that respondents do not use the same decision criteria for public and private goods under the three treatments.

Response Time

Average decision times differ among the 155 pairs at the .0001 level of significance, which supports the obvious conclusion that some choices require more thought and are more difficult than others. This finding may reflect differences in knowledge about the goods in question, differences in experience with the goods in a market exchange context, and differences in the value contrasts within the pairs. For example, choices between public goods and choices between public and private goods



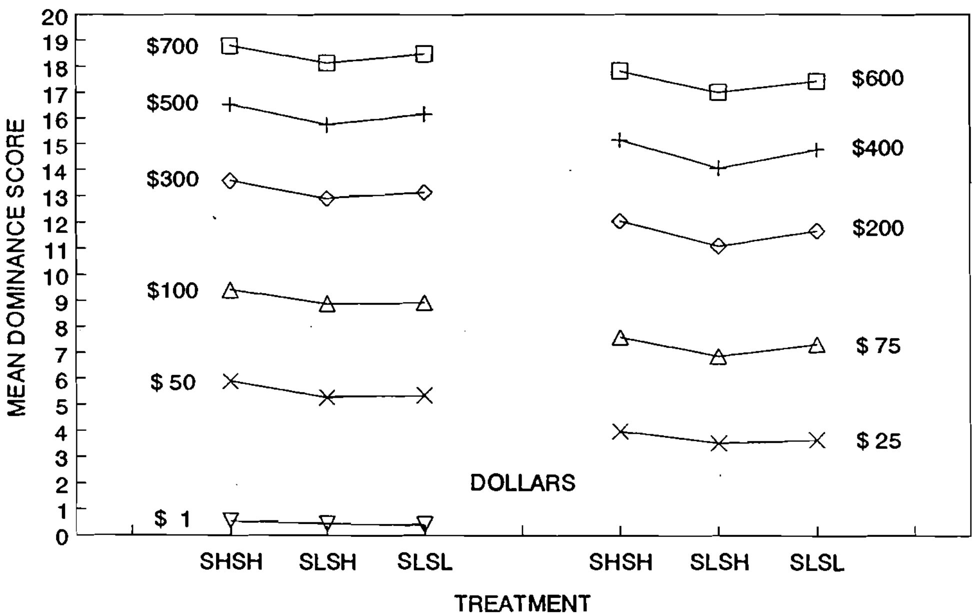


FIGURE 6
MEAN DOMINANCE SCORE PROFILES

require more time on average than choices between private goods or choices involving sums of money. And, because of the range of values involved, choices involving smaller sums of money ($1 to $50) and larger sums of money ($300 to $700) tend to require less time than choices involving mid-range sums ($75 to $200). Figures 7 and 8 illustrate these findings graphically.

Although such general conclusions may seem trivial and self-evident by common sense, they may have significant implications about the validity and reliability of CVM estimates of monetary value for public goods. The time data may also reveal important information about sequential trends as respondents progress through the experiment, or about differences in the way respondents think about different goods or different comparisons. Such questions must await further and more careful analysis of the data.

Classification of Respondents by Preference Type

Earlier pilot studies showed strong nonrandom differences among respondents' preference orderings, thus indicating significantly different response types or market segments that may be explainable and perhaps predictable by such things as demographics, held values (e.g. environmental personality), or political leanings. The paired comparison data allow comparison of preference ordering across individuals, but this, too, must await further analysis.

CLOSURE

The purpose of this paper has been to explore integration of the psychometric method of paired comparisons with neoclassical microeconomic consumer theory to develop a method for estimating WTA for nonmarket goods and services. The objective is to develop a conservative hypothetical market measurement method that avoids loss aversion while also requiring the respondent to consider substitutes and opportunity costs. The proposed method uses paired comparison to order preferences among public goods, private goods, and sums of money, thereby bracketing nonmarket goods between priced private goods and monetary magnitudes. Under appropriate conditions, paired comparison also allows

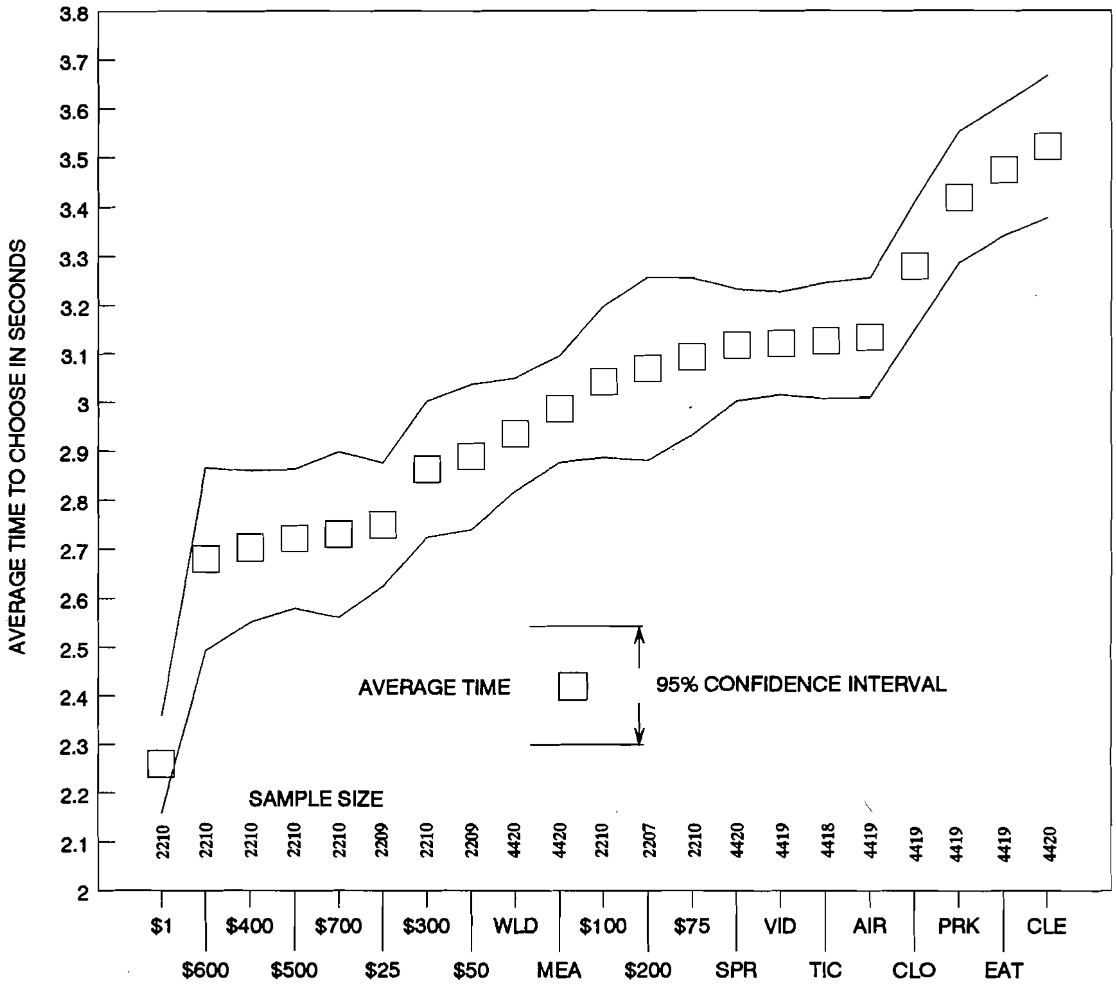


FIGURE 7
AVERAGE TIME TO CHOOSE

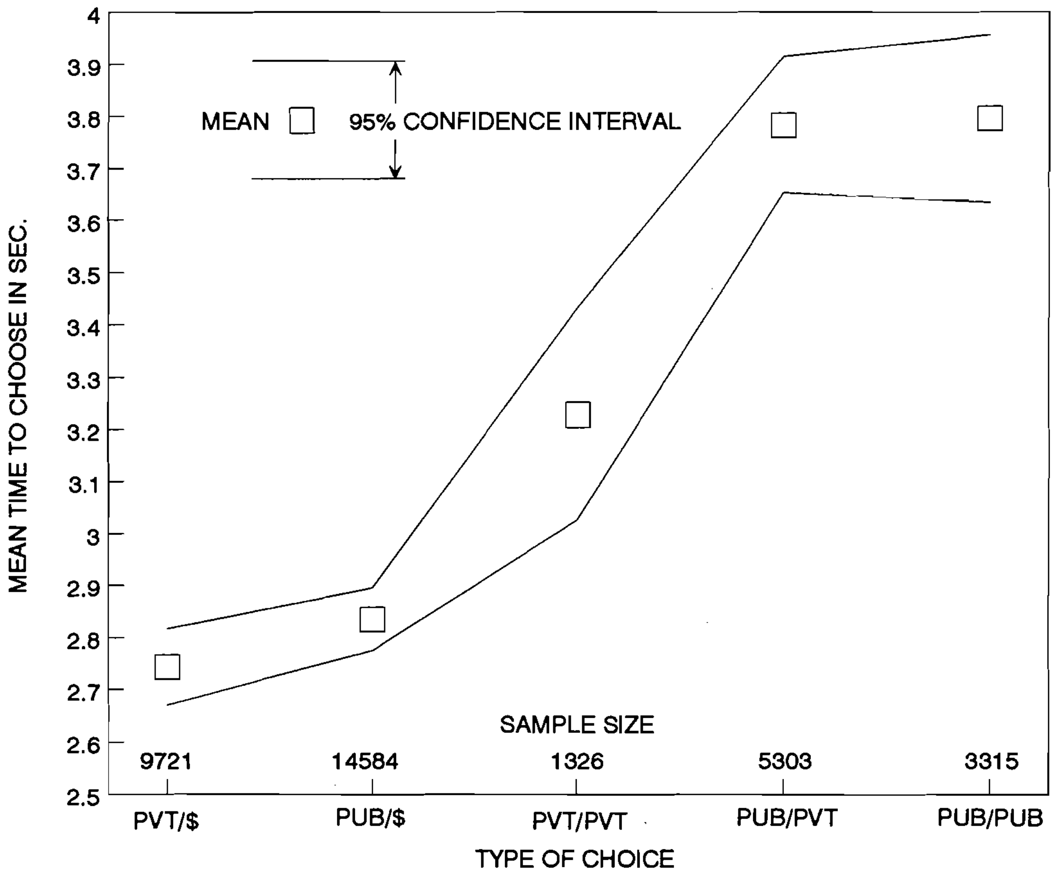


FIGURE 8
AVERAGE TIME TO CHOOSE BY TYPE OF CHOICE

derivation of an interval scale which, in turn, allows interpolation of price estimates for nonmarket goods from bracketing monetary benchmarks.

The empirical section of the paper presents preliminary descriptive results in order to demonstrate application of the method and the richness of information thus obtained. Because of the preliminary nature of the analyses reported herein and the specialized nature of the sample of respondents, the reader must not generalize our results beyond their illustrative purpose. Based on earlier pilot studies and early preliminary results of the present application, we suggest the following hypotheses for further experimental testing:

1. Framing respondents as choosers among alternative gains in an effective and incentive compatible paired comparison experiment ranks the elements of the choice set according to equivalent variation, which is a conservative estimate of willingness to accept compensation (WTA).
2. People tend to choose rationally (with transitivity) among the elements of a diverse set of public goods, private goods, and sums of money in a paired comparison experiment. The cause of intransitivity seems to be either random error for choices too close to call or blatant mistakes that respondents fail to note and correct.
3. Respondents are able to differentiate nonrandomly among the goods and sums of money, thereby achieving a significant ordinal ranking of preferences. This ordinal ranking of preferences describes the individual's utility function. Because the estimated utility function contains both goods and monetary magnitudes, it defines the monetary range within which each good lies
4. Ordinal ranking of preferences varies nonrandomly among respondents, although there is a high percentage of communality. The interpersonal differences are sufficient to require market segmentation before estimating monetary values for nonmarket goods. Such market segmentation can be achieved through standard statistical clustering methods applied to the individuals' dominance profiles.

5. The stimuli in this experiment do not exhibit the degree of homogeneity of variance required for valid aggregate ranking of stimuli and estimation of an interval scale of preference. We must therefore interpret the aggregate ranking and scale calculated in this pilot study as illustrative only. Clustering of the members of a large sample of respondents should yield market segments within which the variance is sufficiently homogeneous to allow valid ranking and scale estimation.

6. With adequate homogeneity of variance, psychometric scaling or linear interpolation may yield more precise monetary estimates for nonmarket goods than is achieved by bracketing on the ordinal ranking of preferences. Or, successive iteration of the experiment with a narrower range of sums of money can more precisely bracket the good(s) of principal interest.

7. Time to choose varies by type of good and type of pair. These differences may be related to knowledge about the goods, experience with the goods in a context of market exchange, and the degree of value difference within pairs. Further analysis also needs to ask whether time to choose varies with the sequence of choice.

8. A rigorous test of the paired comparison method requires a well-developed and communicated information scenario that effectively defines the good(s) in question.

9. With the exception of MEAL, the estimated monetary magnitudes for the goods seem reasonable. Averaged across the three treatments, the method estimates AIR TRAVEL at 71% ($353) of stated retail price ($500), CLOTHES at 93% ($185) of stated retail price, TICKETS at 120% ($90) of stated retail price, and MEAL at 193% ($29) of stated retail price ($15). There may be a tendency to underestimate higher valued goods and overestimate lower valued goods. If so, the indication is that the good(s) of principal interest should be in the mid-range of the values used.

10. The three treatment scenarios (SHARE-SHARE, SOLE-SHARE, SOLE-SOLE) cause some significant differences. These results suggest differences in perceived moral responsibility

and/or differences in decision criteria for public and private goods across the three treatments.

11. The proposed method shows promise for development of standard methods. For example, geologists use an ordinal hardness scale to define and classify the hardness of unknown mineral samples. The scale consists of a set of standard minerals ranked according to ability to scratch and resist scratching by other minerals. An unknown sample is classified on this scale by placing it between adjacent standard minerals such that it scratches one and is scratched by the other. For given market segments, it may be possible to develop standard ordinally ranked sets of goods and or monetary magnitudes such that unknown public goods can be ranked within the standard sets by simple and relatively inexpensive computerized paired comparison experiments.

Bibliography

Arrow, Kenneth J. 1951. Social Choice and Individual Values. New York: Wiley.

Arrow, Kenneth; Robert Solow; Paul R. Portney; Edward E. Leamer; Roy Radner; Howard Schuman. 1993. Report of the NOAA Panel on Contingent Valuation. Federal Register 58(10):4602-4614.

Becker, Gary S. 1965. A Theory of the Allocation of Time. The Economic Journal: The Journal of The Royal Economic Society. Macmillan (Journals) Limited, New York: St. Martin's Press. LXXV(299):493-517.

Bock, R. Darrell; Jones, Lyle V. 1968. The Measurement and Prediction of Judgment and Choice, Robert R. Bush, ed. Holden-Day, Inc., San Francisco, CA. 370 p.

Boulding, Kenneth E. and Sven B. Lundstedt. 1988. Value Concepts and Justifications. In Peterson, George L., B. L. Driver, and Robin Gregory (eds.), Amenity Resource Valuation: Integrating Economics with other Disciplines. State College, PA: Venture.

Cambridge Economics, Inc. 1992. Contingent Valuation: A Critical Assessment. Washington D.C.: Cambridge Economics.

David, H.A. 1988. The Method of Paired Comparisons, Alan Stuart ed. Charles Griffin & Company Limited, London, Great Britain. 188 p.

Deaton, Angus and John Muellbauer. 1981. Economics and Consumer Behavior. New York: Cambridge University Press.

Edwards, Allen L. 1957. Techniques of Attitude Scale Construction. Richard M. Elliott, ed., Kenneth Mac Corquodale, asst. ed. Appleton-Century-Crofts, Inc. New York. 256 p.

Fechner, G. T. 1860. Elemente der Psychophysik. Leipzig: Breitkopf and Härtel.

Fisher, A.; G. H. McClelland; and W. D. Schulze. 1988. Measures of Willingness to Pay versus Willingness to Accept: Evidence, Explanations, and Potential Reconciliation. In Peterson, George L., B. L. Driver, and Robin Gregory (eds.), Amenity Resource Valuation: Integrating Economics with other Disciplines. State College, PA: Venture.

Freeman, A. Myrick III. 1979. The Benefits of Environmental Improvement: Theory and Practice. Baltimore, MD: Johns Hopkins University Press for Resources for the Future.

Gordon, Irene M. and Jack L. Knetsch. 1979. Consumer's Surplus Measures and the Evaluation of Resources. Land Economics 55(1):1-10.

Guilford, J.P. 1954. Psychometric Methods. H.F. Harlow, ed. McGraw-Hill Book Company Inc., New York. 597 p.

Henderson, James M. and Richard E. Quandt. 1980. Microeconomic Theory: A Mathematical Approach. New York: McGraw-Hill.

Hicks, J.R. 1956. A Revision of Demand Theory. Oxford University Press, Amen House, London E.C.4. 196 p.

Kahneman, Daniel; Jack L. Knetsch, and Richard H. Thaler. 1990. Experimental Tests of the Endowment Effect and the Coase Theorem. Journal of Political Economy 98(6):1325-1348.

Kahneman, Daniel., Jack L. Knetsch, and Richard Thaler. 1991. The Endowment Effect, Loss Aversion, and Status Quo Bias. Journal of Economic Perspectives 5(1):193-206.

Kahneman, D. and A. Tversky. 1984. Choices, Values, and Frames. American Psychologist 39:341-350.

Kendall, Maurice; Gibbons, Jean Dickinson. 1990. Rank Correlation Methods, Fifth Edition. Edward Arnold, Hodder & Stoughton Limited, Suffolk. 260 p.

Knetsch, Jack L. 1984. Legal Rules and the Basis for Evaluating Economic Losses. International Review of Law and Economics 4:5-13.

Luce, R. D. 1959. Individual Choice Behavior. New York: Wiley.

Luce, R. D. 1977. The Choice Axiom after Twenty Years. J. of Math. Psych. 15:215-233.

Lancaster, Kelvin J. 1966. A New Approach to Consumer Theory. Journal of Political Economy, LXXIV. 132-157.

Maxwell, A.E. 1974. The Logistic Transformation in the Analysis of Paired-Comparison Data. Br. J. math. statist. Psychol. 27: 62-71.

McFadden, Daniel. 1973. Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), Frontiers in Econometrics. New York: Academic Press.

Morishima, Michio. 1959. The Problem of Intrinsic Complementarity and Separability of Goods. Metroeconomica. 188-202.

Peterson, G. L. and R. D. Worrall. 1970. An Analysis of Individual Preferences for Accessibility to Selected Neighborhood Services. Highway Research Record (NAS--NAE) 305:99-111.

Simon, Herbert A. 1985. Human Nature in Politics: The Dialogue of Psychology with Political Science. The American Political Science Review 79:293-304.

Thurstone, L. L. 1927. A Law of Comparative Judgment. Psychological Review 34:273-286.

Torgerson, Warren S. 1958. Theory and Methods of Scaling. John Wiley & Sons, Inc. New York. 460 p.

Tversky, A. 1969. Intransitivity of Preferences. Psychol. Rev. 76:31-38.

Tversky, A. and D. Kahneman. 1981. The Framing of Decisions and the Psychology of Choice. Science 211:452-458.

Yellott, J. I. 1977. The Relationship between Luce's Choice Axiom, Thurstone's Theory of Comparative Judgement, and the Double Exponential Distribution. J. of Math. Psych. 15:109-144.

A RANDOM UTILITY MODEL APPROACH TO ANALYZING FARM STEWARDSHIP PROGRAMS

by

Joseph Cooper

Russ Keim

Tim Osborn[1]

[1] The authors are in the Resources and Technology Division of the Economic Research Service, U.S. Department of Agriculture, 1301 New York Ave., NW, Washington, DC 20005. The views expressed herein are the authors' and do not necessarily represent the views of the U.S. Department of Agriculture. The authors would like to thank Daniel Piper, ERS, for his valuable assistance.

A RANDOM UTILITY MODEL APPROACH TO ANALYZING FARM STEWARDSHIP PROGRAMS

INTRODUCTION

In response to increasing public concern over the contribution of agricultural pollutants to the degradation of surface and ground water supplies, the 1990 Food, Agriculture, Conservation and Trade Act (FACTA) authorized the USDA to initiate the Water Quality Incentive Program (WQIP). WQIP is administered by the Soil Conservation Service (SCS) through the Agricultural Conservation Program (ACP). Its goal is to mitigate the negative impacts of agricultural activities on ground and surface water supplies through the use of stewardship payments and technical assistance to operators who agree to implement approved practices. With these incentives, farmers are encouraged to experiment with more environmentally benign production practices that they otherwise would not adopt. For most practices the program offers a flat per acre rate with a maximum of $3,500 per contract. In 1992 and 1993 the funding levels for WQIP were $6.75 million and $15 million respectively. Currently, farmers in only a small number of watersheds are eligible to enter the program. However, the issue has been raised (e.g., Sinner, 1990) of making this type of incentive payment program more widely available.

The WQIP incentive payments are not determined through market interaction. Low participation rates in the current eligible areas suggests that the current payments offer insufficient incentive to the farmers to learn about and implement the externality-reducing practices. Given that the WQIP budget is finite, one hundred percent participation is not necessarily possible or cost effective. Our goal is to model the probabilities of participation as a function of a range of incentive payment offers. This response function would be useful in cost-benefit studies comparing the benefits and costs of the various preferred management practices. In addition, since the farmer is free (within limits) to determine how many acres to put into the program at the offered incentive payment level, our second goal is to determine how many acres the farmer will use the new practice on given that the farmer will adopt the practice.

Given the inadequacy of the previously available data, establishing the minimum payment levels to achieve a desired level of participation is not simple. At first glance, if one could determine a farmer's cost of switching production practices, then in theory one could induce the farmer to switch practices by offering a cost share that at least covers the increase in net costs, or if the switch changes revenues per acre as well, a cost share that covers the change in profits (assuming the change in profits is negative).

Unfortunately, it is notoriously difficult to determine the costs of individual components of the production process. Added to this, we would have to predict the changes in output on the targeted land. For example, to use a profit differential approach to determine the appropriate cost share that would induce farmer A to adopt conservation tillage, we would have to know the portion of the farmer's total operating costs devoted to the traditional tillage practice and how this would change if conservation tillage was adopted. Plus, we would need a production function to tell us how output would change on that acreage. To determine these costs and changes in production would be expensive, time consuming, and quite complicated.

Furthermore, even if we could determine the difference in profits for a switch to an alternative practice, an incentive payment based on that differential may not be sufficient to induce the farmer to adopt the practice. For example, the farmer may be a yield, and not a profit, maximizer. The farmer may be risk averse: even if the offered cost share might appear profitable on paper, the farmer may be unwilling to adopt the alternative practice unless the farmer sees neighboring farmers adopting it. Some of the preferred practices may actually reduce farm costs yet have low levels of use. On the other hand, the farmer may be environmentally inclined and if the farmer believes that it will result in an increase in environment benefits, may be willing to switch practices even if the incentive payment does not fully cover the anticipated cost increases. Hence, an incentive payment based on a profit differential may be higher than is necessary to induce the farmer to adopt the practice. To avoid these problems associated with estimating minimum willingness to accept (WTA) to change practices as the difference in cost or profit between the two states, one can use a direct relevation technique for assessing WTA.

THEORETICAL BASIS FOR ESTIMATING WILLING TO ACCEPT IN THE CASE OF NO MARKET TRANSACTIONS

Compared to attempting to estimate a farmer's change in profit under program participation, estimating a farmer's minimum WTA is relatively straightforward and inexpensive. The contingent valuation method (CVM) is a direct elicitation approach used to estimate a respondent's willingness to accept or pay for a nonmarket good. CVM is a market simulation approach in which the survey respondent is provided with scenario involving a hypothetical market for good and is asked to value the good. Since CVM is frequently used to value public goods that the respondent is unfamiliar with, such as wildlife preservation, the survey responses can be very sensitive to the formulation of the hypothetical market in the survey instrument. Unlike most nonmarket goods, because the farming practices to be discussed here are hardly hypothetical to the farmer, it is relatively straightforward to apply CVM to determine the incentives needed to adopt these practices.

The earliest CV questions simply asked respondents to state their maximum WTP or minimum WTA for, example, the change in the level of access to some environmental amenity. However, one can now choose among several elicitation formats. Of several possible formats for CVM, referendum, or dichotomous choice, CVM (DC CVM) is particularly suitable for developing a model that predicts minimum WTA for nonadopters. Under this approach, the respondent is prompted to provide a "Yes" or "No" response to a dollar bid amount contained in the valuation question. The bid amount is be varied across the respondents. This method is particularly likely to reveal accurate statements of value since the format provides reasonable incentives for value formulation and reliable value statement (Hoehn and Randall, 1987).[2] In fact, the National Oceanic and Atmospheric Administration Blue Ribbon Panel's (co-chaired by Kenneth Arrow and Richard Solow) proposed guidelines for conducting natural resource damage assessment using CVM suggest that all CV studies should use the referendum format (U.S.

---

[2] While willingness to pay (WTP) questions are considered to be incentive compatible in the referendum format, some capacity for strategic response bias (in both the upper and lower directions) may still exist with WTA questions. However, the referendum format most likely diminishes this bias over the open-ended question format.

Department of Commerce, 1993). With DC CVM, instead of trying to identify the farmer's profit function (which, at any rate, would not include any profit-independent reasons to accept the program), we simply need to determine whether or not the farmer's minimum WTA is greater than or equal to the offered payment incentive.

The farmer's decision process is modelled using the random utility model approach. From the utility theoretic standpoint, a farmer is willing to accept \$C to switch to a new production practice if the farmer's utility with the new practice and incentive payment is at least as great as at the initial state, i.e., if $U(0,y;x) \leq U(1,y + C;x)$, where 0 is the base state; 1 is the state with the WQIP practice; y is farmer i's income; and x is a vector of other attributes of the farmer that may affect the WTA decision. C can be written as $C^* + \delta$, where δ is state 0 pecuniary costs less state 1 pecuniary costs. Hence, C can be considered a 'net' incentive payment. Note that δ can be positive; due to some nonpecuniary costs, a farmer may not have switched to the preferred practice even if δ is positive. The farmer's utility function U(i,y;s) is unknown due to components of it that are unobservable to the researcher, and thus, can be considered a random variable from the researcher's standpoint. The observable portion is V(i,y;x), the mean of the random variable U. With the addition of an error ε^i, where ε^i is an i.i.d. random variable with zero mean, the farmer's decision to accept \$C can be re-expressed as

[1] $V(0,y;x) + \varepsilon^0 \leq V(1,y + C;x) + \varepsilon^1.$

If $V(i,y;x) = \gamma^i + \alpha y$, where $\alpha > 0$, for i = 0,1, then the farmer is willing to accept \$C for the change if $\gamma^0 + \alpha y + \varepsilon^0 \leq \gamma^1 + \alpha(y+C) + \varepsilon^1$.

The decision to accept \$C can be expressed in a probability framework as $\Pr\{WTA \geq \$C\} = \Pr\{V^0 + \varepsilon^0 \leq V^1 + \varepsilon^1\} = \Pr\{\varepsilon^0 - \varepsilon^1 \leq V^1 - V^0\}$, where $V^1 - V^0 = \gamma + \alpha C$, and where $\gamma = \gamma^1 - \gamma^0$. Since $V^1 - V^0 = \gamma + \alpha C$ is generated directly from the utility model given above, it is compatible with the theory of utility maximization. The mean of the random variable WTA can be expressed as

[2] $$E(WTA) = \int_0^\infty [1 - F_\varepsilon(\Delta V)]dC - \int_{-\infty}^0 F_\varepsilon(\Delta V)dC.$$

where $F_\varepsilon(.)$ is the cumulative density function. Because the utility difference function can be expressed in this probability framework, the logit or probit qualitative dependent variable regression models can be used to estimate the coefficients.

Once the parameter estimates have been obtained, mean WTA can be calculated numerically using the formula for the mean of a random variable. However, if the CDF in (2) is logistic or normal, then a closed form solution to E(WTA) is available as follows (Hanemann, 1989):

[3] $$E(WTA) = \frac{\gamma}{\alpha}.$$

or equivalently, decomposing γ into the sum of its parts in a multiple regression case (with p explanatory variables) yields

[3.1] $$E(WTA|x_0) = \frac{(\gamma_1 x_{0.1} + \gamma_2 x_{0.2} + \ldots + \gamma_p x_{0p})}{\alpha},$$

where x_0 is a vector of the explanatory variables (excluding the bid variable) evaluated at, say, their means. For a symmetric distribution, this mean value is also equal to the median value. For this paper, the mean, or median value is of secondary importance to estimating the probability of participation in the program for a schedule of incentive payments. These can simply be obtained through $P_i = F_\varepsilon(\Delta_i)$. From a cost effectiveness standpoint, the optimal rates of acceptance may not be the same for each practice.

EMPIRICAL BASIS FOR THE DETERMINING MINIMUM WILLINGNESS TO ACCEPT AND THE LEVEL OF PARTICIPATION IN THE WQIP

Traditionally in the CVM literature, univariate probit or logit is used to analyze the responses. However, the referendum data may not be nonrandomly selected from the survey respondents as only those respondents who do not currently use the preferred practices were asked the referendum questions.

Regressing the referendum data without accounting for the nonrandom selection of this data from the survey data set can produce biased and inefficient coefficient estimates. For the survey, a sample selection question was used to identify respondents who do not currently use the practice. Next, respondents who said that they did not currently use the practice were asked the WTA question. Formally, denoting the 0/1 response to the sample selection question as y_{2i} and denoting the 0/1 response to the WTA for participation question as y_{1i}, y_{1i} is observed only when $y_{2i} = 0$. In other words, the disturbances are correlated between the two questions. The system of equations is presented in utility difference form as

[4.1] $\Delta V_{1i} = x_{1i}'\gamma_1 + \alpha C_i = \varepsilon_{1i}$ where $y_{1i} = 1$ if $\Delta V_{1i} \leq \varepsilon_{1i}$, $y_{1i} = 0$, otherwise,

[4.2] $\Delta V_{2i} = x_{2i}'\gamma_2 = \varepsilon_{2i}$ where $y_{2i} = 1$ if $\Delta V_{2i} \leq \varepsilon_{2i}$, $y_{2i} = 0$, otherwise

$[\varepsilon_{1i}, \varepsilon_{2i}] \sim$ bivariate normal $(0,0,1,1,\rho)$,

where 4.1 is the WTA equation discussed in the previous section, $y_{1i} = 1$ if farmer i's true WTA is greater than the bid offer, $\Delta V_{1i} = V^0_i - V^1_i$, $x_i^{1\prime}\gamma = x^0_{1i}{}'\gamma_0 - x^1_{1i}{}'\gamma_1$, $\varepsilon_{1i} = \varepsilon^1_{1i} - \varepsilon^0_{1i}$, and C is the incentive payment offer. Using the same format as (4.1), Equation (4.2) is the sample selection equation. Assuming a bivariate normal relationship for ε_{1i} and ε_{2i}, bivariate probit is used to estimate the two sets of coefficients. The bivariate probit with sample selection log-likelihood function for the situation where y_{1i} is observed only when $y_{i2} = 0$ is:

[5] $$lnL(\gamma_1,\gamma_2,\rho_{12}) = \sum_{y_{2i}=0, y_{1i}=1} \ln\phi_a[x_{1i}'\gamma_1, -x_{2i}'\gamma_2, -\rho_{12}]$$

$$+ \sum_{y_{2i}=0, y_{1i}=0} \ln\phi_a[-x_{1i}'\gamma_1, -x_{2i}'\gamma_2, \rho_{12}] + \sum_{y_{2i}=1} \ln\phi[x_{2i}'\gamma_2]$$

where C is included in X_1 for notational simplicity, Φ is the normal CDF, Φ_a is the bivariate CDF, and ρ_{12} is the correlation coefficient between the two equations. Because the likelihood function in equation

(5) contains more information than would a univariate probit likelihood function for equation (4.1), maximization of equation (5) offers efficiency gains over univariate probit. Furthermore, equation (5) accounts for potential correlation between (4.1) and (4.2) and therefore corrects for the sample selection bias that could occur if (4.1) were to be estimated singly (Boyles, Hoffman, and Low, 1989). The disadvantages of the bivariate log-likelihood function in equation (5) is that convergence of the estimates is not always easily achieved with it and estimated covariance matrices are frequently singular. Note that if estimated $\rho_{12} = 0$, then the farmers who answer the WTA question can be assumed to be randomly drawn from the sample and equation 4.2 can be ignored. Equation (4.1) can then be estimated using probit.

Applying the definition of conditional probability, the farmer response function for the bivariate probit case is

[6] $\text{Prob}(\ WTA_i \geq bid_i \mid y_{2i} = 0) = \Phi_a(\mathbf{x}_{1i}'\gamma_1, -\mathbf{x}_{2i}'\gamma_2, \rho)/\Phi(-\mathbf{x}_{2i}'\gamma_2)$,

where Φ_a is the bivariate normal probability and Φ is the normal probability. Of course, if $\rho_{12} = 0$, then $\text{Prob}(\ WTA_i \geq bid_i \mid y_{2i} = 0) = \Phi(\mathbf{x}_{1i}'\gamma_1)$.

One may hypothesize a priori that potential explanatory variables for $\mathbf{x}_{1i}$ and $\mathbf{x}_{2i}$ include whether or not the farmer believes the practice will affect farm profitability, soil type, type of crop(s) planted, total farm size, amount of training needed to implement the practice, and level of environmental awareness and concern, etc. Except for the bid offer, both equations could use the same explanatory variables.

As stated earlier, estimating the probit or bivariate probit with sample selection models is the first step of our research agenda. In addition to developing the farmer participation equation as a function of the offer amount, we would also like to know the amount of acres the farmer will enroll given the decision to participate. The number of acres enrolled in the preferred practice by farmer *i* can be stated as

$$[7]\ \mathrm{PACRES}_i = z_1'\theta + u_i,$$

where PACRES_i is the amount of acres in the preferred practice, z_1 is a vector of explanatory variables, and u_i is a disturbance with mean zero. Explanatory variables can include the payment offer, length of participation in the program, total acreage, level of environmental concern, erosion potential, and type of crop to be grown on the eligible acres.

Unfortunately, ordinary least squares estimates of equation (7) on farmers who do not currently use the preferred practice but agree to do so with the incentive payment are potentially biased. Because these hypothetical acreage enrollments are only observed for the farmers who answered "yes" to the WTA question, the sample for equation (7) is not drawn randomly from the population who answered the survey, implying omitted variable bias. Furthermore, additional bias may be added as only those answering "no" to the sample selection question asked the WTA question. In addition to being potentially biased, OLS estimation of equation (7) is inefficient (Greene, 1990). The estimate of equation (7) can be corrected by considering the responses to the two qualitative dependent variable questions in the analysis of (7). Dubin and McFaddens' (1984) or Heckman's two step procedure, or 'Heckit' (Heckman, 1979) have been used for estimation of systems consisting of one qualitative dependent variable and one continuous equation.

In this paper, an extension of the Heckit procedure to three equations is used for estimation (Tunali, 1986; Greene, 1992) when ρ is statistically different from 0. Since PACRES_i is observed only when $y_{1i} = 1$ and $y_{2i} = 0$, the revised version of equation (7) is

[8] E[PACRES$_i$ | z_i, in sample] =

$$= \text{E}[\text{PACRES}_i \mid z_i,\ y_{1i} = 1,\ y_{2i} = 0\ ]$$

$$= \text{E}[\text{PACRES}_i \mid z_i,\ \varepsilon_{1i} \geq \Delta V_{1i},\ \varepsilon_{2i} < \Delta V_{2i}\ ]$$

$$= z_i'\theta + \text{E}[u_i \mid \varepsilon_{1i} \geq x_{1i}'\gamma_1 + \alpha C_i,\ \varepsilon_{2i} < x_{2i}'\gamma_2]$$

Tunali (1986) shows that equation (8) reduces to

[9] PACRES$_i$ $= z_i'\theta + \lambda_{1i}\tau_1 + \lambda_{2i}\tau_2 + \eta_i,$

where η_i is a disturbance term. λ_{1i} and λ_{2i} are defined as:

[9.1] $$\lambda_{1i} = \phi(-x_{1i}'\gamma_1)\Phi[(-x_{2i}'\gamma_2 - \rho_{12}y_{1i})/(1 - \rho_{12}^2)^{1/2}]/\Phi_a$$

$$\lambda_{2i} = \phi(-x_{2i}'\gamma_2)\Phi[(-x_{1i}'\gamma_1 - \rho_{12}y_{2i})/(1 - \rho_{12}^2)^{1/2}]/\Phi_a,$$

where $x_1 = [x_1, C]$ and $\beta_1 = [\gamma, \alpha]$ and where Φ_a = bivariate normal CDF $\Phi(x_1'\beta_1,-x_2'\beta_2,-\rho_{12})$. The derivatives $\partial\lambda_{1i}/\partial$Bid and $\partial\lambda_{2i}/\partial$Bid are less than zero. Failure to include λ_1 and λ_2 in the regression may lead to a biased estimate of the vector of coefficients θ. Greene (1992) describes the formulation of (7) in more detail as well as providing the asymptotic covariance matrix for it. If ρ_{12} is not statistically different from zero, then one does not need to be concerned with the potential sample selection bias dues to omission of those who currently use the preferred practice. Instead, we need to only be concerned with potential bias in the estimation of (6) due to inclusion of only those farmers who answered 'yes' to the WTA question. In this case, the basic Heckit model is used:

[10] PACRES$_i$ $= z_i'\theta + \lambda_i\tau + \eta_i,$

where η_i is a disturbance term. λ is defined as $\lambda_{1i} = \phi(x_{1i}'\gamma_1)/\Phi(x_{1i}'\gamma_1)$, where γ_1 is the vector of probit coefficients from the WTA question and where $\partial\lambda_{1i}/\partial$Bid < 0. When using the correct asymptotic covariance matrix (see Greene, 1981; Heckman, 1979) in the linear regression of PACRES on z and λ, the coefficient estimates are consistent, if not efficient (though more efficient than OLS).

Because the survey sampled some regions at higher rates than others (e.g. noncropland areas were sampled at lower rates than cropland areas), the data were scaled by sampling weights. Multiplying the data by the weights gives greater weight to the observations from the regions with the lower probability of being selected and decreases the weight to the observations from the regions with higher probability of being selected. For estimation, the weights are multiplied by the sample size divided by the sum of the weights so that the sum of the weights across the observations is the sample size (Greene, 1992).

DATA DESCRIPTION

The Area Studies project is a data collection and modelling effort directly involving the Economic Research Service (ERS), the U.S. Geological Survey (USGS) and the National Agricultural Statistical Service (NASS). Data on cropping and tillage practices and input management were obtained from a comprehensive field and farm level survey of about 1000 farmers in each of four critical watershed regions: the Eastern Iowa and Illinois Basin areas, the Albermarle-Pamlico Drainage Area covering Virginia and North Carolina, the Georgia-Florida Coastal Plain and the Upper Snake River Basin Area. These study areas were selected from within the set of U.S. Geological Survey's National Water Quality Assessment (NAWQA) sites and sample sites were chosen to correspond to Soil Conservation Service's National Resource Inventory (NRI) so that information on the physical characteristics corresponding to farming activities would be available. For example, slope and erosion potential of the soil would seem to be factors that may influence the decision to adopt conservation tillage.

Information about the extent of the farmers' current use of the preferred practices as well as their willingness to adopt these practices if they do not currently use the practice were provided by a supplemental questionnaire. Respondents to the comprehensive questionnaire were asked to complete and mail in this additional section. For the final analysis, 1261 observations were available.

The practices analyzed here, a short description of each, and the current incentive payment levels are:

Conservation Tillage (CONTIL) - Tillage system in which at least 30% of the soil surface is covered by plant residue after planting to reduce soil erosion by water; or where soil erosion by wind is the primary concern, at least 1,000 pounds per acre of flat small grain residue-equivalent are on the surface during the critical erosion period. Incentive payment not to exceed $12 per acre.

Integrated Pest Management (IPM) - Pest control strategy based on the determination of an economic threshold that indicates when a pest population is approaching the level at which control measures are necessary to prevent a decline in net returns. This can include scouting, biological controls and cultural controls. Incentive payment not to exceed $12 per acre.

Legume Crediting (LEGSR) - Involves estimating the amount of nitrogen available for crops from previous legumes (e.g. alfalfa,clover, cover crops, etc.) and reducing the application rate of commercial fertilizers accordingly. Incentive payment not to exceed $10 per acre for row crops.

Manure Testing (MANTST) - Estimating the amount of nutrients available for crops from applying livestock or poultry manure and reducing the application rate of commercial fertilizer accordingly. Incentive payment not to exceed $10 per acre for row crops.

Split Applications of Nitrogen (SPHN) - Applying one-half or less of required amount of nitrogen for crop production at or before planting, with the remainder applied after emergence, in order to supply nutrients more evenly and at times when the crop can most efficiently use them. Incentive payment not to exceed $10 per acre for row crops.

Soil Moisture Testing (SMTST) - Use of tensiometers or water table monitoring wells to estimate the amount of water available from subsurface sources. Incentive payment not to exceed $10 per acre.

All of these practices are currently being supported by WQIP. For the willingness to adopt question for all of the practices except conservation tillage the bids offered are ($2, $4, $7, $10, $15, $20). For conservation tillage the bids are ($4, $6, $9, $12, $18 and $24). The bid ranges were chosen to cover what we perceived to be the likely range of WTA. The bids were randomly assigned with equal probability to the surveys.[3] The specific referendum CVM question asked to the farmer is "If you don't use this practice [listed in the question] currently, would you adopt the practice if you were given a $[X] payment per acre?" (answer 'yes' or 'no'). The sample selection equation is "Is this practice [listed in the survey] currently in use on your farm?" (answer 'yes' or 'no').

The pool of variables from which the explanatory variables were drawn is:

TACRE - Total acres operated.

EDUC - Formal education of operator.

EINDEX - Sheet and rill erosion index.

FLVALUE - Estimated market value per acre of land.

EXPER - Farm operator's years of experience.

SNT - Soil nitrogen test performed in 1992 (dummy).

TISTST - Tissue test performed in 1992 (dummy).

CTILL - Conservation tillage used in 1992 (dummy).

PESTM - Destroy crop residues for host free zones (dum).

ANIMAL - Farm type-beef,hogs,sheep (dummy).

GRAINS - Farm type-cash grains (dummy).

ROTATE - Grasses and legumes in rotation (dummy).

MANURE - Manure applied to field (dummy).

HEL - Highly erodible land (dummy).

ESTIMATION RESULTS

Table 1.1 and 1.2 present the weighted univariate and bivariate probit results for the sample selection question (i.e., the question of whether or not the farmer currently uses the practice; Table 1.2) and for the WTA for adoption question (Table 1.1). Convergence of the bivariate model was achieved

[3] The survey procedures in place did not allow a more complex allocation of bids. See Cooper (1993) and Kanninen (1993) for other possible surveys designs.

for IPM and MANTST. For the other practices, univariate probit was used for the estimation of each equation. For bivariate normal densities (though not necessarily for other densities), a ρ_{12} of 0 would imply that the two equations are independent. If significant, a negative correlation is expected as y_1 can equal 1 only if $y_2 = 0$. Of the six practices, the correlation coefficient between the two equations (ρ) is significantly different from 0 only for IPM. Hence, univariate probit is sufficient for the estimation when ρ equals zero, as is the case for MANTST.[4]

With regards to the other coefficients, in Table 1.1, the key variable, BID, is of the correct sign and is significant to at least the 1% level for four of the practices and is significant at the 5% level for one of the practices. BID does not achieve the 5% level of significance for IPM. In general, explanatory power among the other variables was lower. Of the other variables common to all six practices, the coefficient on years of education was significant and positive for there of the six practices in Table 1.1 while coefficients on years of experience were negative and significant for 3 of the practices. Years of experience was not significant for any practice in the results in Table 1.2. To test for regional differences in the responses, regressions were tried with dummies for the regions but none of the associated coefficients were significant. The difficulty in observing variables that actually factor into the farmer's decision on whether or not to adopt the practice demonstrates the benefits of the stated preferences approach used here over an indirect approach, such as one that relies on estimating a profit function.

Using the bivariate probit coefficient estimates from Tables 1.1 and 1.2 for IPM, and univariate probit results from Table 1.2 for the other practices, Table 2 presents the median WTA per acre estimates for each of the six practices. The median for IPM in the bivariate probit case was determined numerically for the offer amount at which the conditional probability in equation (6) equals 50%. Weighted univariate probit estimates range from \$32 per acre for CONTIL to \$57 per acre for

[4] The univariate probit results are available upon request from the authors.

MANTST.[5] For five of the six practices, the unweighted univariate results were higher than the weighted results. Standard errors for the univariate medians are calculated using an analytic approach described by Cameron (1991) and are presented in parentheses below each of the values. The weighted bivariate median for IPM is two dollars higher than weighted univariate median. The WTA results in Table 2 show that if 50% participation is desired of the current nonusers, the required payments would have to be much larger than the current WQIP payments.

For the continuous equations modelling the acres enrolled given the decision to participate, Table 3 presents the Heckit regression results for the practices where ρ was insignificant and the bivariate probit with sample selection regression results where it was significantly different from zero (the IPM case). In the table, for the Heckit results, λ_1 is the coefficient on the Mills ratio for the WTA question. For the bivariate probit case for IPM, λ_1 and λ_2 are the coefficients for the variables defined in equation (9.1). For λ_1 and λ_2 for IPM, the coefficient on the former is significant. Hence, $\partial EACRES/\partial BID$ cannot be isolated considered in isolation of λ_1. However, for the Heckit regressions λ_1 is insignificant and $\partial EACRES/\partial BID$ can be estimated in isolation of λ_1.

Not too surprisingly, of the other coefficients, total acres enrolled (TOTACRE) tended to be the best predictor of acres enrolled and was highly significant for all the programs. Interestingly, BID was significant to at least 5 percent for only IPM and MANTST, which seems to that for the other practices that although the bid level may be a factor in the decision to participate in the program, physical constraints, and not price, may determine how many acres are enrolled. The coefficient on FLVALUE was positive and significant to at least the 5% level for two of the practices and to at least the 1% level for two others.

[5] For the curious, the weighted univariate grand means (γ) are -0.6275, -1.0966, -1.211, -1.5421, -1.52341, and -1.48393, respectively for the practices in Table 2.

MODEL APPLICATIONS

Given that the WTA estimates necessary to encourage 50% of current nonusers to adopt (Table 2) are much higher than the current payments levels, it is not surprising that participation in the program by eligible farmers is quite low for many of the practices. However, given that encouraging participation is not costless, a cost-efficiency or cost-benefit analysis could be used to determine what participation rates, and hence what offer amounts would be desirable for each practice. To do this, a farmer response function is necessary. As discussed earlier, the probit coefficient results can be plugged into the normal CDF to predict probability of adoption of the practices for different incentive payment levels. In conjunction, for those farmers who are predicted to adopt the practice at a given payment level, the continuous equation can be used to predict the number of acres enrolled.

Using the univariate probit coefficients results for the WTA equation, Figure 1 presents graphs of the relationship between the offer amount and the probability of acceptance for those farmers who do not currently use the practices. All the adoption function show some positive adoption rate with an incentive payment of $0. In particular, the adoption function for CONTILL predicts the highest adoption rate (27%) with a $0 incentive payment. Interestingly, practicing conservation tillage may actually reduce per acre farming costs (Skinner, 1990; see *Reduced Tillage*, Table 1). In fact, CONTILL is also the most popular of the examined preferred practices, with over 70% of farmers in the data set currently using the practice.[6] The positive adoption rate at $0 suggests that some current nonusers may be willing to adopt the practice without an incentive payment provided that they are given sufficient information on the practice.

Figure 2 presents the univariate and bivariate response functions for IPM. The bivariate response function is a representative farmer's probability of adoption of the desired practice given that the farmer does not currently use that practice. The Figure 2 shows that for IPM, univariate probit

[6] The percentage of farmers in the data set currently using each practice is 1 minus the ratio of observations (1) and observations (2) listed on the bottom of Table 1.2.

underpredicts adoption rates for incentive payments below the median and overpredicts adoption for incentive payments above the median.

The appendix describes a model that could use these response functions to determine the incentive payments that maximize the net benefits of the incentive payment program, where net benefits are defined as the change in environmental benefits (in dollars) due to the switch to the preferred practices minus the total incentive payment outlays. Further research is needed to put a monetary value on the environmental benefits of the changes in farm management practices.

SUMMARY

Farmers can be encouraged to voluntarily adopt environmentally sound management practices through the use of incentive payments. Current USDA practice is to offer a fixed "take it or leave it" payment per acre to those not currently using the desired practices. Hence, there is insufficient observed data to model the probability of farmer adoption of the environmentally sound management practices as a function of the payment offer. Without this function, one does not know what at what level to set incentive payments to achieve desired levels of participation. This paper uses a direct relevation technique based on a random utility model to develop and estimate models predicting farmer adoption of the practices as a function of the payment offer. Models that predict the acreage enrolled given the decision to accept the incentive payments are also developed and estimated. These results can be used in a cost-benefit analysis to best decide how to allocate the program budget among the preferred production practices.

REFERENCES

Boyes, W., Hoffman, and S. Low. 1989. "An Econometric Analysis of the Bank Credit Scoring Problem." Journal of Econometrics 40:3-14.

Cameron, T.A. 1991. "Interval Estimates of Non-market Resource Values from Referendum Contingent Valuation Surveys." Land Economics 67(4):413-21.

Joseph Cooper. 1993. "Optimal Bid Selection for Dichotomous Choice Contingent Valuation Surveys," Journal of Environmental Economics and Management, Vol. 24:25-40.

Dubin, J. and D. McFadden. "An Econometric Model of Residential Appliance Holdings and Consumption." Econometrica 52:345-362.

Esseks, J. and S. Kraft. 1988. "Why Eligible Landowners Did Not Participate in the First Four Sign-ups of the Conservation Reserve Program." Journal of Water and Soil Conservation 43 (May-June):251-256.

Greene, W. 1992. Limdep: User's Manual and Reference Guide, Version 6.0. Econometric Software, Inc., Bellport, New York.

Greene, W. 1990. Econometric Analysis. MacMillan Publishing Company, New York, NY.

Hanemann, M. 1989. "Welfare Evaluations in Contingent Valuation Experiments with Discrete Response Data: Reply." American Journal of Agricultural Economics 71(4):1057-1061.

Hanemann, M. 1984. "Welfare Evaluations in Contingent Valuation Experiments with Discrete Response Data." American Journal of Agricultural Economics, 66(3):332-341.

Heckman, J. 1979. "Sample Selection Bias as a Specification Error." Econometrica 47:153-161.

Hoehn, J. and A. Randall. 1987. "A satisfactory benefit cost estimator from contingent valuation." Journal of Environmental Economics and Management, 12:226-247.

Kanninen, B. 1993. "Optimal Experimental Design for Double-Bounded Dichotomous Choice Contingent Valuation." Land Economics 69(2):138-46.

Osborn, T., F. Llacuna, and M. Linsenbigler. 1992. "The Conservation Reserve Program: Enrollment Statistics for Signup Periods 1-11 and Fiscal Years 1990-92." USDA/ERS, Statistical Bulletin No. 843.

Sinner, J. 1990. "We Can Get More for Our Tax Dollars." Choices 5(2):10-13.

Tunali, I. 1986. Research in Labor Economic 8:235-282.

U.S. Department of Commerce, National Oceanic and Atmospheric Administration. 1993. "Proposed Rules: Natural Resource Damage Assessment." Federal Register, 58(10): 4601-4614. 15 CFR, Chapter IX.

APPENDIX

The results presented in this paper may be applied in the following simplified model for the USDA to choose optimal incentive payments ($PYMT_j$) such that environmental benefits minus program costs are maximized.

$$[1] \qquad MAX \qquad \Sigma_i(\Sigma_j B_{ij} * EACRES_{ij}) - \Sigma_i(\Sigma_j PYMT_j * EACRES_{ij})$$

subject to

$$[2] \qquad EACRES_{ij} \leq TACRES_i \quad for \ all \ i,j$$

$$[3] \qquad \Sigma_j(\Sigma_i EACRES_{ij} * PYMT_j) \leq BUDGET$$

$$[4] \qquad \Sigma_j(EACRES_{ij} * PYMT_j) \leq FLIMIT \quad for \ all \ i$$

$$[5] \qquad PYMT_j \leq PYMT_j^{max} \quad for \ all \ j$$

$$[6] \qquad PYMT_j \geq 0$$

where i = 1,...,n farms, j = 1,...,m preferred practices and the expected number of acres enrolled

$$[7] \qquad EACRES_{ij} = P(PYMT_j)*f_{yes}(TACRES_i,PYMT_j) + [1- P(PYMT_j)]*f_{no}(TACRES_i,PYMT_j),$$

where P(.) is the farmer i's probability of acceptance and is based on the probit results and f(.) is the number of acres enrolled given the decision to participate ($PACRE_{ij}$) and is based on the sample selection model results. Note that $[1-P(PYMT_j)]*f_{no}(TACRES_i,PYMT_j)$ drops out of the expected value function in (7) since we assume that farmers who do not accept the payment will not put any acres in the preferred practice. We assume that for administrative reasons, $PYMT_j$, which is the incentive payment for practice $_j$, varies only across the practices. $EACRES_{ij}$ is the number of acres each farmer enrolls in practice j. BUDGET is the level of available funding for the program. $TACRES_i$ is the total size of farm i or it may be some other upper limit, such as eligible acreage, on available acres per farmer and could be unique for each i,j. FLIMIT is the maximum total payments per farm. $PYMT_j^{max}$ is the maximum possible payment per acre due to some administrative or political reason. B_{ij} is the dollar value of the environmental benefits of the new practice over the traditional practice. Note that B_{ij} varies by farm as some farms may show bigger environmental gains by switching practices than others.

Equation 1 is the environmental benefits ($) minus the USDA's incentive payment outlays. Equation 3 places a ceiling on the total cost of the program. Equation 4 is the per farm limit on payments. Note that the shadow price associated with equation 3 is especially interesting for policy analysis, such as benefit-cost comparisons. This shadow price is the change in net environment benefits ($) for an incremental change in the budget.

Table 1.1. Weighted Univariate and Bivariate Probit Results: Model for the Decision to Participate in Each of the Programs By Farmers Not Currently Using the the Practice.

	PROGRAM					
	CONTILL	SPHN	IPM	LEGSR	MANTST	SMTST
Variable			Coefficient Estimates			
CONST	−0.2863	−0.6686 **	−1.1790 **	−1.8774 **	−1.5626 **	−1.5574 **
	(−0.94)	(−3.07)	(−4.32)	(−7.90)	(−7.48)	(−7.67)
BIDVAL	−0.0198 *	0.0433 **	0.0154	0.0303 **	0.0266 **	0.0324 **
	(1.94)	(5.11)	(1.45)	(3.33)	(3.34)	(4.16)
EDUC	−0.0666	−0.0308	0.2562 **	0.1002 *	0.0512	0.0959 **
	(−1.23)	(−0.74)	(7.94)	(2.29)	(1.34)	(2.54)
CTILL	0.0906	−−	−−	−−	−−	−−
	(0.51)					
HEL	−0.1139	−−	−−	−−	−−	−−
	(−0.58)					
TISTST	−−	0.8757 *	−−	0.2365	−0.2527	−−
		(2.10)		(0.77)	(−0.73)	
EXPER	−0.0079	−0.0133 **	−0.0022	−0.0004	−0.0105 **	−0.0096 *
	(−1.39)	(−2.97)	(−0.68)	(−0.09)	(−2.43)	(−2.21)
PESTM	0.1869	−−	0.4538 **	−−	−−	−−
	(1.13)		(4.40)			
ROTATE	0.0306	0.0356	0.0789	0.4422 *	0.2868	0.2034
	(0.10)	(0.16)	(0.44)	(1.69)	(1.42)	(0.97)
MANURE	−0.0202	−0.2206	−−	−0.3039	0.2468 *	−−
	(−0.09)	(−1.43)		(−1.57)	(1.92)	
ANIMAL	−−	0.0742	−−	0.2942 *	0.3979 **	−−
		(0.57)		(2.13)	(3.55)	

– Coefficient divided by standard error in parentheses.
– * = significance of 5%. ** = significance of 1%.

Table 1.2. Univariate and Bivariate Probit Results – Continued: Model for Whether or Not Farmer Currently Uses the Practice.

	PROGRAM					
	CONTILL	SPHN	IPM	LEGSR	MANTST	SMTST
Variable	Coefficient Estimates					
CONST	0.4894 **	−0.6355 **	−1.3434 **	−1.0057 **	−1.9535	−1.8991 **
	(2.87)	(−4.33)	(−8.47)	(−6.59)	(−8.47)	(−8.38)
EDUC	−0.0075	0.0911 **	0.2321 **	0.1377 **	0.0914	0.0911 *
	(−0.23)	(3.15)	(7.70)	(4.64)	(2.08)	(2.10)
CTILL	0.8769 **	– –	– –	– –	– –	– –
	(9.83)					
HEL	0.0631	– –	– –	– –	– –	– –
	(0.57)					
TISTST	– –	0.7551 **	– –	0.0690	−0.1741	– –
		(3.34)		(0.31)	(−0.47)	
EXPER	0.0005	0.0010	−0.0028	−0.0010	0.0023	0.0035
	(0.16)	(0.33)	(−0.85)	(−0.29)	(0.49)	(0.74)
PESTM	−0.3947 **	– –	0.4544 **	– –	– –	– –
	(−3.68)		(4.41)			
ROTATE	0.0734	−0.2081	0.0729	0.4319 **	0.2031	−0.0116
	(0.39)	(−1.15)	(0.39)	(2.60)	(0.91)	(−0.04)
MANURE	0.0306	−0.0318	– –	0.3602 **	0.4248	– –
	(0.25)	(−0.29)		(3.38)	(3.12)	
ANIMAL	– –	−0.2362 **	– –	0.0675	0.2031	– –
		(−2.43)		(0.71)	(1.59)	
RHO	NA	NA	−0.9849	NA	−0.9783	NA
			(−3.22)		(−0.00)	
Obs. (1)	331	683	830	860	1101	1070
Obs. (2)	1243	1198	1202	1204	1192	1186

– * = significance of 5%. ** = significance of 1%.

Table 2. Median Minimum Expected Willingness to Accept (Per Acre) To Encourage Use of the Practices.

A. Univariate Probit

CONTIL	SPHN	IPM	LEGSR	MANTST	SMTST
$29.55	$27.39	$45.10	$56.92	$61.63	$51.37
(6.75)	(4.11)	(11.16)	(16.16)	(17.13)	(12.07)

B. Weighted Univariate Probit

CONTIL	SPHN	IPM	LEGSR	MANTST	SMTST
$31.67	$25.35	$42.06	$50.96	$56.59	$45.76
(10.38)	(3.14)	(9.39)	(12.18)	(13.76)	(8.51)

C. Weighted Bivariate Probit

CONTIL	SPHN	IPM	LEGSR	MANTST	SMTST
--	--	$44.00	--	--	--

– Standard errors are presented in parentheses.

Table 3. Weighted Sample Selection Model for Acres Enrolled Given the Decision to Participate in the Program (Dependent Variable = Acres Enrolled).

	PROGRAM					
	CONTILL	SPHN	IPM	LEGSR	MANTST	SMTST
Variable	Coefficient Estimates					
CONST	225	−298	−2376 **	411	−112	−163
	(0.34)	(−1.02)	(−3.15)	(0.55)	(−0.21)	(−0.17)
TACRE	0.05 **	0.158 **	0.728 **	0.220 **	0.103 *	0.396 **
	(2.51)	(5.63)	(48.20)	(4.53)	(2.26)	(9.67)
FLVALUE	−0.06	0.109 *	0.131 **	0.163 **	0.092	0.104 *
	(−1.13)	(2.27)	(4.66)	(2.38)	(1.56)	(1.80)
BIDVAL	−7.59	3.933	39.864 *	−0.032	19.108 *	8.267
	(−0.77)	(0.61)	(2.26)	(−0.00)	(2.23)	(0.71)
EDUC	43.86	42.594 *	179.870 **	−20.388	8.171	−3.910
	(1.24)	(2.13)	(3.48)	(−0.55)	(0.24)	(−0.07)
SNT	−−	−−	−−	33.802	−5.289	−−
				(0.287)	(−0.040)	
PESTM	220.51 *	−−	−−	−−	−−	−−
	(1.89)					
ANIMAL	−243.23 *	−22.16	−181.60 **	−231.23 *	−60.78	−365.51 **
	(−1.85)	(−0.22)	(−3.91)	(−1.91)	(−0.43)	(−3.22)
GRAINS	−44.59	71.031	−14.099	−81.509	−28.290	−11.857
	(−0.46)	(0.77)	(−0.31)	(−0.82)	(−0.25)	(−0.13)
LAM−1	−44.59	168.88	646.27 **	−188.82	−21.76	126.77
	(0.08)	(1.00)	(3.05)	(−0.57)	(−0.10)	(0.28)
LAM−2	NA	NA	207.80	NA	NA	NA
			(0.92)			
Obs.	73	123	115	64	78	100
R−Sq.	0.23	0.26	0.77	0.30	0.17	0.56

− * = significance of 5%. ** = significance of 1%.

Figure 1. Response Curves for the Subsidized Practices
(Results based on the univariate normal distribution)

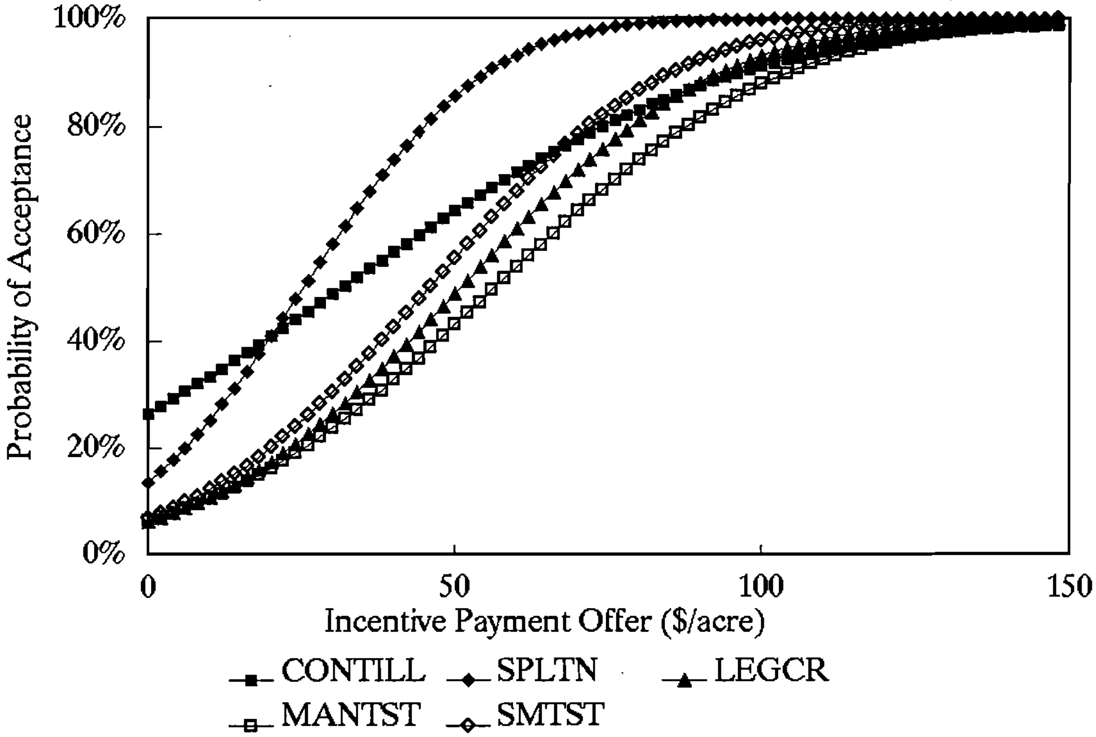
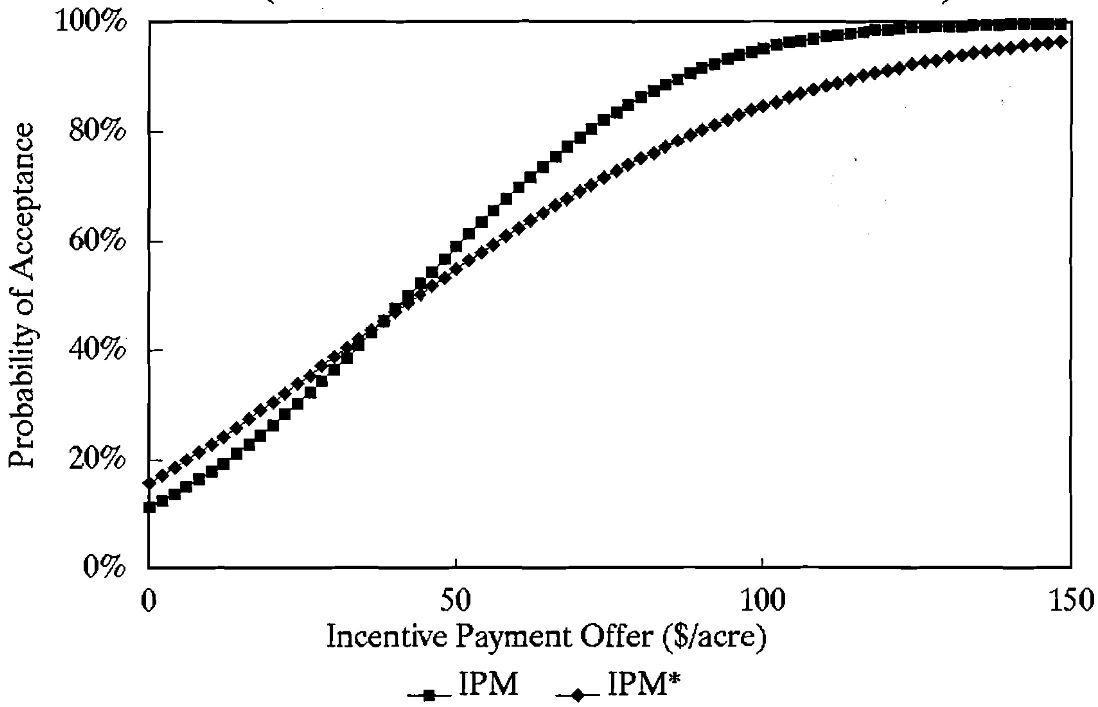


Figure 2. Response Curves for Integrated Pest Management
(Curves for univariate and bivariate normal results)



IPM* is based on the conditional bivariate normal distribution.

EXPOSURE-BASED RISK PERCEPTIONS AND CONDITIONAL DAMAGES AND BENEFITS*

Gregory L. Poe
Department of Agricultural, Resource, and Managerial Economics
Cornell University

and

Richard C. Bishop
Department of Agricultural Economics
University of Wisconsin-Madison

ABSTRACT

This paper examines how subjective perceptions of the exposure-risk transformation function are affected by baseline exposure levels, and discusses the implications for estimating damage and benefit functions. Using nitrates found in individual wells, evidence is provided that perceptions of health risks at each exposure level are affected by baseline exposure levels. A function of conditional damages and benefits is estimated from contingent valuation data, with marginal damages and benefits reaching a peak at an intermediate level of nitrates and then declining. Possible explanations for this non-convexity are provided.

* This research and paper were funded in part by the Center for Integrated Agricultural Systems, University of Wisconsin-Madison, and the College of Agriculture and Life Sciences, Cornell University. The authors are indebted to Jean-Paul Chavas, Bill Provencher, and Richard Boisvert for their comments on this research, and to Patricia Champ and Dan Mullarkey for help with the data collection.

EXPOSURE-BASED RISK PERCEPTIONS AND CONDITIONAL DAMAGES AND BENEFITS

Damage and benefit functions that link monetary values to the concentrations of pollutants are fundamental to environmental economics. Broadly defined, these functions measure the economic loss or gain across exposure levels subject to the condition that reference levels and utility are held constant. The conventional damage function approach adopts a zero or low pollution state as a common reference point, while the benefits approach holds reference conditions at a level that is currently experienced or accessible [Freeman]. Implicit in the identification of a single reference level is the assignment of property rights, as this level serves as a basis for determining welfare losses and gains.

With groundwater contamination and other environmental risks that are individually experienced and for which property rights are not clearly established, a commonly held reference level is inconsistent with the potential Pareto improvement criterion. Here, the relevant reference condition for welfare measurement is the level of exposure and expected utility experienced by individuals at the time of the policy determination. To the extent that subjective perceptions of health risks associated with a specific exposure level depend on the reference level of risk, willingness to pay (WTP) for changes in exposure will be conditional upon reference exposure levels. Instead of a single damage or benefits function, an alternative conditional damages approach envisions a conditional damage function associated with each initial level of exposure.

Using nitrate levels found in individual well water and contingent valuation, this paper estimates how conditional WTP values for a 25 percent reduction in exposure levels and for avoiding a 25 percent increase in exposures are affected by initial nitrate levels.

EXPOSURE-BASED RISK PERCEPTIONS AND CONDITIONAL DAMAGE AND BENEFIT FUNCTIONS

With respect to exposure to nitrates in well water (N), the consumer's choice problem can be characterized by the minimization of the *ex ante* planned expenditure function [Smith]

$$\dot{e}(g(h;N,S(N)),p,q,N,\overline{EU}_N) = \min_{S(N),x} p'[N,S(N)]+q'x \quad \text{subject to } EU = \overline{EU} \tag{1}$$

where: $\dot{e}(\cdot)$ is the planned expenditure function; $g(h;N,S(N))$ is the subjective distribution of health outcomes (h) given nitrate exposure levels (N) in personal well water and averting consumption of water (S(N)) from alternative sources such as bottled water, water from home filtration systems, and water transported from a 'pure' source; p is the corresponding state independent vector of prices for different water sources including explicit or implicit prices for water drawn from private wells; q is the state independent vector of prices for all other goods (X); and $\overline{EU}_N$ is the expected utility referenced by nitrate level. Under these conditions, this formulation is the dual of the option price model suggested in Crocker, Forster and Shogren, with the exception that nitrate levels are directly incorporated into the expenditure difference function here in order to capture 'non-use' motivations.

Willingness to pay (WTP) for a 1-δ percent reduction in nitrate exposure is given by

$$WTP_\delta = \dot{e}(g(h;N,S(N)),p,q,N,\overline{EU}_N) - \dot{e}(g(h;\delta N,S(\delta N)),p,q,\delta N,\overline{EU}_N) \tag{2}$$

Similarly, WTP for a project that avoids a certain increase in contamination of τ -1 is specified as

$$WTP_\tau = \dot{e}(g(h;\tau N,S(\tau N)),p,q,\tau N,\overline{EU}_{\tau N}) - \dot{e}(g(h;N,S(N)),p,q,N,\overline{EU}_{\tau N}) \tag{3}$$

Corresponding to the direction of the change in nitrate exposure, WTP_δ and WTP_τ are referred to as benefit and damage measures respectively. Because the reference condition is the without project status, WTP is a compensating measure in both models. At the limit, as 1-δ and τ approach 1, these incremental measures become marginal WTP values.

It is important to recognize that the representations in Equations 2 and 3 are different than the 'standard' presentation of damage and benefit functions: in the standard approaches the expected utility value is held at some common reference level (say the utility level associated with N^o) that may or may not be related to the initial or target exposure levels associated with the proposed change. For example, if the nitrate reference point was $N^{\circ}= 2$ mg/l, then the WTP for a 1-δ reduction from an arbitrary level N would be given as

$$\mathbf{WTP_{\delta} = \acute{e}(g(h;N,S(N)),p,q,N,\overline{EU}_2) - \acute{e}(g(h;\delta N,S(\delta N)),p,q,N,\overline{EU}_2)} \tag{4}$$

If it is assumed that state dependent utilities in the healthy and unhealthy states are the same regardless of exposure levels and identical across consumers, then any lack of correspondence in the risk-income trade-offs measured in equations (2) and (4) would be attributable to differences in risk perceptions and expected utility indices. Focusing on the former, recognition that risk perceptions are subjective and conditional upon reference levels is a feature of prospect theory [Kahneman and Tversky] and prospect reference theory [Viscusi]. In a WTP framework, the effect of exposure-based, or conditional, subjective probability perceptions on risk-income trade-offs is characterized in Figure 1. In this figure, the expected utility loci across income and objective risks are dependent upon subjective conditional risks: that is $g(P|P^M)$ is the subjective risk function conditional upon the "objective" reference risk level P^M, and $g(P|P^U)$ is the subjective risk perception associated with an initial "objective" risk level P^U. As depicted, both marginal and total WTP values are assumed to be affected by their baseline levels of risk.

One implication of exposure-based subjective probabilities and divergent expected utility loci is that a stepwise pattern of measuring the value of reducing objective risks from P^u to P^L that sequentially aggregates AB and B'C' will lead to a biased estimate of the value of a complete reduction measured by AC. To our knowledge, a study by Römer and Pommerehne [1990] of hazardous waste contamination in Germany provides the only such "path" comparisons of WTP for sequential health risk reductions. As depicted by a summary of their results in Table 1, WTP for a risk reduction from 0.0001 to 0.00005 was significantly higher for individuals with an initial reference risk level of 0.0005 when compared with those with a reference risk level of 0.0001. While such differences may be attributed, in part, to income effects [Römer and Pommerehne] or anomalies of the contingent valuation method such as sequencing bias, and embedding or "warm glow" effects [Mitchell and Carson;

Kahneman and Knetsch], the results depicted in Table 1 are not inconsistent with the concept of exposure-based subjective risk perceptions.

The implication from the Römer and Pommerehne study and the hypothesis of exposure-based subjective risk perceptions is that marginal or incremental WTP across exposure levels cannot be pooled to form a single total damages function. Instead, conditional damage or benefit functions need to be estimated for each distinct reference set of exposures. This result is analogous to path dependency in the welfare theory of price changes [Just, Hueth, and Schmitz]. In this case, however, the path dependency is attributed to different perceptions of risk associated with reference exposure levels, rather than differences in utility indices.

NON-CONVEXITIES IN DAMAGE AND BENEFIT FUNCTIONS

Irrespective of the reference level used, little consensus has emerged in the theoretical and empirical economics literature concerning the convexity of WTP relationships across risk levels. Early theoretical models from the statistical life literature supported the maintained hypothesis that marginal WTP should be convex and rise with the level of risks [Jones-Lee; Weinstein, Shepard and Pliskin]. More recent formulations indicate that convexity of WTP for risk reductions is indeterminate when averting behavior is possible unless restrictive *a priori* assumptions are imposed [Shogren and Crocker; Quiggen]. Empirical evaluations of WTP to reduce hypothetical risks have similarly found conflicting evidence. In a study of transportation risk Jones-Lee *et al.* [1985] provided evidence supporting convexity of damages across risk levels. In contrast, Smith and Desvousges [1987] observed increasing marginal valuation with decreasing baseline risks in a study of risk reductions associated with a toxic waste disposal site[1].

The studies mentioned so far focused on the relationship between WTP and risk. More conventionally, damage functions are expressed as a function of physical exposure levels rather than

[1] Evidence of non-convexities in environmental damages for "technical" reasons have been discussed by Repetto [1981, 1987]. This source of non-convexity is, however, beyond the scope of this paper.

health risk [e.g. Conrad and Olsen; Xepapadeas]. Convexity in damages, or lack thereof, across exposure levels may be determined by the convexity of the transformation function between exposure levels and health risks. Even if damages across probabilities were convex, a sufficiently concave transformation function between exposure levels and subjective health risks could result in a convex damage function across exposure levels.[2] For instance, an individual's subjective transformation function may be such that a sudden jump from a zero health risk to a positive probability occurs at a positive non-zero "threshold" level of exposure [Kask and Maani]. In other words, below a certain level of exposure (T) individuals simply assume that risk is zero so that $g(h;N)= 0 \ \forall \ N \leq T$.[3,4] Some evidence of threshold effects are found in radon studies, which demonstrate that subjective risks anchor on government action and safety limits of exposure and shift dramatically at these points [Smith *et al.*]. Discontinuities of this sort may induce an unexpectedly large WTP for small shifts in objective risks that cross subjective threshold levels simply because the change in perceived probability of adverse health effects far exceeds the shift in objective probability.

It is also arguable that discontinuities may be more pervasive than the single threshold model, in that individuals combine continuous ranges of probabilities into discrete safety groupings such as definitely safe, probably safe, etc. Such judgmental heuristics could imply a step damage function: marginal damages of a shift in exposure would conceivably be zero if the increment in exposure did not

---

[2] Page and Ferejohn [1974] raise this issue in discussing the convexity of "environmental transfer functions" for production externalities.

[3] In an influential paper, Lichtenstein *et al.* [1978] observed that there is a tendency to overestimate low probability events such that there is a discontinuity between subjective risks and observed frequencies at zero. The approach offered here suggests that this discontinuity may actual occur at a positive level of "objective" risks. See Kask and Maani [1992] for further discussion.

.[4] The function g(N) is used here rather than g (N, S(N) to indicate the subjective health probabilities in the absence averting actions.

involve a shift to another safety level [i.e $g(h;\delta N)=g(h;N)$], and the magnitude of damages might be large if a small shift in exposure involves a perceived change in safety levels.[5]

The convexity of damages might be also be affected by averting opportunities [Zeckhauser and Fisher; Shitaba and Winrich; Repetto, 1987; Shogren and Crocker]. First, if averting actions offer complete protection and the cost of averting is unaffected by the level of contaminant to be removed (e.g. bottled water), then marginal damages associated with increased ambient levels could conceivably be zero once individuals adopt averting actions. If complete averting behavior is adopted, the effective exposure and risk remains constant regardless of contamination level. Anecdotal evidence from this research supports the hypothesis that endogenous averting actions that are already undertaken are viewed as substitutes and act to lower WTP. In response to a $216 dichotomous choice bid value for reducing exposure levels, one respondent wrote, "No, but I would have if I hadn't recently put in a H20 softener and reverse osmosis system for this reason". A pre-survey participant indicated that his WTP was bounded because he was able to "truck" all the good quality drinking water from his daughter's well in town. With high investment in transporting equipment, this alternative represented a relatively permanent solution.

To the extent that the probability of adopting such averting behavior is positively correlated with ambient risks [Smith and Desvousges, 1985], the aggregate damage function would be concave from below at high levels of exposure as a greater proportion of households adopt effective averting practices. Moreover, WTP may be bounded simply by the opportunity for substitution. As for any commodity with reasonable substitutes, a choke price will likely exist above which the commodity is not consumed. Thus, we would expect marginal damages to be bounded due to opportunities for substitution.

Combined, the subjective perceptions of risk, non-linearities in the exposure-safety transformation function, and averting opportunities suggest that damage and benefit functions will have

---

[5] Kopp and Smith [1993, p. 128-29] use a similar argument in discussing the relationship between dollar damages and an index of physical injury associated with oil spills.

both convexities and non-convexities. A plausible depiction of total and marginal damage functions are depicted in Figure 2. In this figure the sharp increase in WTP is associated with crossing of threshold levels, and total WTP is truncated by averting opportunities. By similar logic, it is expected that functions of conditional damages will have local concavities and convexities.

SURVEY DESIGN

This study was conducted in rural portions of Portage County, Wisconsin which do not have municipally provided water (N_{1990} = 22,432). Portage County has had extensive nitrate contamination problems in the last two decades, and past research suggests that 18 percent of private wells exceed government standards for nitrate (NO_3-N) of 10 mg/l. The source of elevated nitrates in this region is attributed to agricultural activities upgradient from wells [Portage County Groundwater Management Plan].

A sequential two-stage survey design was used to measure nitrate exposure levels and to elicit contingent values. Households participating in the survey were randomly drawn from a private mailing list covering the targeted rural areas that did not have municipal water supplies. In the first stage (Stage 1), individuals were asked to complete an initial questionnaire and to submit water samples that would be tested at the Wisconsin State Laboratory of Hygiene for nitrates. In the second stage (Stage 2), the participants who had returned the Stage 1 questionnaire and a water sample were provided their nitrate test results, general information about nitrates, and a graphical depiction of their exposure levels relative to maximum natural levels of nitrates and government safety standards (see Appendix). Thus, when answering the Stage 2 contingent valuation questions used in this analysis, individuals had a full set of specific information focusing on their current exposure levels and general information about sources of nitrates, possible health risks, government standards for nitrates, and possible mitigating activities. Remedial options for individual households included repairing or improving the existing well, constructing a new well, purchasing bottled water, and installing a denitrification system. Participants were informed that filtration systems would cost about $200 to $420 per year and that bottled water would cost $480 to $720 for a three member household.

The implementation of the survey followed established procedures detailed in Dillman. A total of 480 Stage 1 surveys were mailed. After correcting for bad addresses (n_{BAD}= 39) the response rate to both stages was approximately 64 percent. Nitrate levels ranged from not detectable to over 43 mg/l with a mean of 5.90 mg/l. Approximately 16 percent of the tests exceeded government standards of 10 mg/l.

EVIDENCE OF EXPOSURE-BASED EXPOSURE-SAFETY TRANSFORMATION FUNCTIONS

General safety perceptions reflect government health standards for nitrates of 10 mg/l. Using a return potential response format, participants were asked "*Suppose that your water test had indicated one of the nitrate levels listed below. In **your opinion** would you believe that this well is safe or unsafe for **your household** to use as the primary source of drinking and cooking water?*". Nitrate levels included 2, 4, 6, 8, 10, 12, 15, 20, 30 and 40 mg/l and response categories were "*Definitely Safe*", "*Probably Safe*", "*Probably not Safe*", and "*Definitely not safe*". Figure 3 provides average safety responses for the three different reference groups corresponding to levels at or below natural nitrate levels found in Wisconsin aquifers (low: <2 mg/l), levels corresponding to evidence of human impacts but within nitrate health standards (moderate: 2-10 mg/l), and levels that exceed nitrate safety standards (high: >10 mg/l). In all cases the rate of decrease in safety perceptions is highest across the 8-12 mg/l range, which suggests that risk perceptions are anchoring on government health standards. Yet, at the same time it is clear that the standard is not necessarily regarded as a safe/unsafe threshold: some individuals appear to accept the threshold while others consider it to be too conservative or liberal. Thus, the safety standard is not perceived to demarcate safety zones by all respondents, which conflicts with the safe/unsafe default assumptions used in past research of groundwater valuation [Edwards; Sun, Bergstrom and Dorfman].

Importantly, the distribution in safety perceptions differs in a systematic manner that is consistent with an exposure-based risk perception hypothesis. Participants with "low" nitrate levels within natural bounds perceive "moderate" exposure levels to be relatively unsafe when compared to participants in the with "moderate" or "high" reference nitrate levels. At the other extreme, respondents experiencing "high" exposure levels are relatively tolerant of "moderate" and "high" exposure levels than the other groups[6].

In all, the evidence from this survey indicates that subjective risk/safety perceptions are indeed influenced by reference exposure levels. Individuals with different baseline exposure levels have different perceptions of the exposure-risk transformation function. This result supports the rationale for pursuing conditional damage and benefit functions.

ECONOMETRIC MODEL OF WTP

Two separate contingent valuation questions regarding incremental changes in exposure were posed to each respondent. One dichotomous choice question asked respondents to consider a 25 percent decrease in their exposure levels using the following format to convey target and reference levels of exposure:

* *With the groundwater protection program the nitrate levels in **all** Portage County wells would be reduced by **25 percent** over the next five years. This means that the nitrate levels in your well would fall to ___________ mg/l.*

* *Without the groundwater protection program, please assume that the nitrate levels in Portage County will remain at their **current level**. This means that without the program your nitrate level will remain at ___________mg/l.*

This question, which corresponds to WTP_δ in equation (2) above, is referred to as the "incremental benefits" function in the analyses that follow, reflecting the fact that lower exposure levels are expected

[6] The following comparisons of low, medium and high levels were significantly different at the 5 percent level using a difference of means test: low-medium, NO_3=2,4,6 mg/l; low-high, NO_3=2,4,6,8,10,12,15 mg/l; medium-high, NO_3= 6,8,10,12,15,20 mg/l.

to reduce, or at least not increase, the subjective probability of illness. The term incremental is used because the magnitude of the change being evaluated.

Target and reference levels for the dichotomous choice question that elicited WTP to avoid an increase in exposure (incremental damages) were depicted as follows:

* *__With__ the groundwater protection program, nitrate levels in all Portage County wells __will definitely__ be kept at their current levels. This means that the nitrate levels in your well will remain at _________mg/l, and that future increases in nitrate levels would be avoided.*

* *__Without__ the groundwater protection program, please assume that it has been estimated that the nitrate levels in all Portage Count wells __would rise by 25 percent__ over the next five years. This means that the nitrate levels in your well would rise to ________ mg/l if this groundwater program is not adopted.*

In both the benefit and damage formats the reference and target exposure levels were individually inscribed in the blank spaces of each survey.

Following these descriptions of the reference and target conditions, a YES/NO response was elicited for the following contingent valuation question.

Would you vote for the groundwater protection program described above is the total <u>annual</u> cost to your household (in increased taxes, lower profits, higher costs, and higher prices) were $___________ each <u>year</u> beginning now and for as long as you live in Portage County?

Dollar 'bid' values were individually inscribed in each survey and ranged from $1 to $999.

With the dichotomous choice format used in this survey, individual values are not directly observed. A YES or NO response to a dichotomous choice question merely provides an indication between bid values (A) and the individual's "true" value, defined here to be WTP. Thus, although an indicator of the valuation is observed, the actual value remains a random variable. Assuming a logistic functional form, the WTP distribution can be estimated by

$$\Pi(\text{YES};X) = \Pi(\text{WTP} \geq A) = 1 - F(A;X'\beta,k) = [1+\exp^{(A-X'\beta)/k}]^{-1} \tag{5}$$

where X is a vector of covariates, β is a corresponding vector of coefficients, and k is a scale parameter [Cameron]. In this formulation it follows that $E(WTP|X)=X'\beta$, where the vector X includes a constant. If X includes nitrate exposure levels, then a function of conditional damages can be estimated for the incremental valuation questions described above. In the analysis that follows polynomial functions of nitrate levels are estimated.

As discussed in the conceptual framework, averting actions should affect the marginal rate of substitution between ambient exposure levels and all other goods. The binary variable DAVTPERM pools two averting actions that are regarded as having high and relatively irreversible investment costs: installing a purification system and trucking water in from another source. As discussed, these actions represent an investment in personal protection, and would be expected to have a negative effect on WTP for marginal risk reductions. A similar conclusion does not necessarily follow for purchasing bottled water as an averting activity. While, from the perspective of substitution this may exert a negative effect on WTP, the fact that some individuals would no longer purchase bottled water if their risk was reduced might also have a positive effect on WTP if bottled water expenditures were relatively large. Given these countervailing influences, there is no sign expectation for the coefficient on the binary averting behavior variable DBOTWAT.

Evidence from the psychological literature suggests that the formulation of risk and safety perceptions is more profound than a simple mapping between exposure and probabilities of adverse health states. Investigating the multiple dimensions of perceived risk, Slovic, Lichtenstein and Fischhoff found that nitrogen fertilizers had high scores along principal "undesirable" dimensions in factor space analyses, and, by extension, have low acceptability [Slovic *et al.*]. Translated into the welfare framework, the perceived benefits of reducing ambient levels are high and should be correlated with certain underlying socio-psychological factors.

Beneficiality: Acceptability of risks have been linked to the voluntariness with which they are incurred [Starr; Slovic] or the perceived beneficiality of exposure [Vlek and Stallen]. For example, in examining air pollution standards, Baird [1986] found that "smelter employed respondents were much more likely to be tolerant of the [air pollution] risk" (p. 432). Due to the strong sample association between perceptions that "*nitrates are a problem in Portage County*" and the belief that "*agricultural fertilizers* *are a major source*" of nitrate contamination,[7] along with the scientific evidence that elevated nitrates in Portage County are linked to agricultural activities [Portage County Groundwater Management Plan], a binary variable DFARM for household involvement in farming was included in the model. The coefficient on this variable is expected to be negative.

Familiarity: People who have lived with exposure over a period of time without observing health effects are less likely to be concerned about environmental hazards. Support for this supposition is provided in previous empirical studies of environmental risk, which have found "years in home" [radon; Smith and Johnson], "number of years living in (town)" [groundwater: Schultz and Lindsay], and "long-time" residency [groundwater: Hamilton] to be negatively correlated with risk perceptions, WTP for environmental protection, and environmental concerns, respectively. To account for this factor, a categorical variable (LIVEPAST) of responses to the question "*About how long have you lived in Portage County*" is included in the analysis, with an expected negative coefficient.

Environmental and Non-Use Motives: Similar to familiarity, environmental concerns have been linked to risk intolerance and WTP for groundwater protection. With respect to groundwater, Edwards [1988] found bequest motives to be a strong contributor to WTP for groundwater protection. Mitchell and Carson [1989] and McClelland *et al.*[1992] have also found strong bequest, stewardship and intrinsic motives for the perceived benefits of groundwater protection. To capture these non-use values, a simple sum (NON-USE) of categorical responses with regards to health concerns of "*future generations*" and "*other people living today*" was created.

Demographic Characteristics: A number of demographic or socioeconomic characteristics have been linked to risk perceptions and contingent values, and are thus incorporated as control variables in the analysis. Adoption of averting behavior [toxic wastes, Smith and Desvousges, 1986], learning about risks [radon: Smith *et al.*], and WTP for risk reductions [transportation: Jones-Lee, Hammerton and

[7] Kruskal's Gamma (γ) statistic for measuring association in ordered variables had a highly positive and significant value of association [0.471(sd=0.058)] between beliefs that "*agricultural fertilizers are a major source of contamination*" and the perception that "*nitrates are a problem in Portage County*" [Poe, p. 108].

Philips] have been found to be negatively correlated with age. The sex of respondent is also routinely included in studies of environmental concerns. While Hamilton [1985] observed a "motherhood effect" in which women with small children viewed water pollution as a particularly serious problem and Viscusi, Magat and Huber [1987] note a "parental altruism" effect in risk-dollar tradeoffs, a general survey of environmental risk studies suggests that the sex of the respondent is not related to environmental concerns [Van Liere and Dunlap]. Various forms of education variables are also a mainstay of risk analyses, with some evidence that there is a negative relationship between education and risk tolerance (e.g. Loomis and Duvair). These parameters are included in the analysis with the variables AGE, DSEX, and DCOLLEGEGRAD.

Definitions, descriptive statistics, and the expected signs of the coefficients are provided for each variable in Table 2.

INCREMENTAL CONDITIONAL DAMAGES AND BENEFITS

The econometric evaluation of a function of incremental damages across nitrate levels was estimated using the logistic function presented in Equation (5) and the two stage-estimation process detailed in Cameron [1988]. The paucity of observations at high nitrate levels and the polynomial approach used to model responses across nitrate levels limited the analysis to an upper bound of 25 mg/l. A lower bound of 0.20 mg/l was set to account for reductions that would be measurable: the testing method used by the Wisconsin State Laboratory of Hygiene was not able to measure nitrate levels below 0.15 mg/l.[8] As a result, 28 observations were deleted from the lower tail and 7 observations were deleted from the upper tail of the nitrate distribution. Because of item non-response, there was a different number of observations for the benefit (n=221) and damage (n=218) estimates.

[8] Further support for this lower truncation point is that the United States Geological Survey assumes that levels less than 0.2 mg/l represent natural background levels. In the USGS classification system, levels from 0.2 mg/l to 3 mg/l are transitional, and may or may not represent human influence. As noted in the nitrate information sheet, 2 mg/l is regarded as the upper bound for natural levels in Wisconsin.

After accounting for this truncation, the number of observations represented about 50 percent of the mailable Stage 1 surveys for both questions.

Econometric estimates for "Full" and "Nitrates Only" forms of the first, second and third order polynomials are provided in Tables 3a and 3b. In general, the significant coefficients on the non-nitrate covariates have the expected sign. Non-use motivations and the age of the respondent have positive and negative coefficients, respectively. Involvement in farming has a negative effect on WTP in the benefits estimate, but is not a significant explanatory variable in the damages model. In contrast, the coefficient on the education variable is positive in the damages estimate, but is not significant in the benefit models. Being a female has a negative effect on WTP in this data set.

It is interesting to note that purchasing bottled water in the past has a positive effect on WTP to avoid a 25% increase in exposure. However, a similar result is not observed for risk reductions. Permanent averting actions were not a significant explanatory variable in either model, a result that might be attributed, in part, to the fact that only a small number of participants had adopted permanent averting actions. Regardless of cause, the results of this analysis do not support the hypothesis that actual averting actions negatively impact WTP.

In spite of this observation, WTP pay for incremental reductions and avoiding incremental increases in exposure levels does appear to be bounded. While incremental WTP rises with exposure levels as indicated by the positive coefficient on the linear models for both the benefit and damage functions, the quadratic and the cubic models suggest that the functions of conditional benefits and damages have regions of convexity. The cubic form, which provides the best 'fit' of the polynomial functions investigated[9], suggests that damages are convex for relatively low levels of exposure, but are eventually concave. Points of inflection are determined to be approximately 7.5 mg/l and 8.0 mg/l for the benefit and damage functions, respectively.

[9] Higher order polynomials were investigated, but added little to the fit of the models. Moreover, the inclusion of higher order polynomials acted to model individual observations at the upper end of the nitrate spectrum.

Using the "Cubic Nitrates Only" estimates, a graphical depiction of WTP for incremental changes in exposure is provided in Figure 4. It is obvious from this figure that WTP reaches a maximum and then diminishes for both benefits and damages.[10] It is interesting to note that the maximum of the benefits function, at 14.9 mg/l, closely corresponds to a point where a 25 percent reduction will place the final level very close to the 10 mg/l standard. The incremental damages function reaches a maximum at 15.9 mg/l. This correspondence between maximum WTP and observed threshold responses is more obvious when incremental values are converted to marginal values. As depicted in Figure 5[11], the maximum values peak at 10.5 for the benefit function and 11.3 for the damage function. Importantly, this analysis of marginal WTP demonstrates that responses are not simply governed by a symbolic or warm glow effect in which WTP values are not affected by level of exposure. If such lack of responsiveness was the case, marginal values derived from an incremental analysis would be a declining function across all nitrate levels.

As noted, the eventual decline in incremental WTP does not appear to be attributed to averting actions that have been undertaken. Instead, it is hypothesized that WTP is bounded simply by the opportunity for substitution through the establishment of a choke price: maximum WTP values of $394/year for incremental benefits and $325/year for incremental damages fall in the range of annual least cost averting expenditures of $200 to $420 for an average household. The fact that incremental WTP for risk reductions actually declines after an intermediate level of exposure may be related to the fact that individuals may not perceive a change in safety levels across high nitrate levels. Over two thirds of the respondents felt that water with high nitrate levels of 15 mg/l or higher was definitely not safe for their household to use as their principal source of drinking water. For those individuals, for

[10] In constructing this figure, WTP was restricted to be non-negative. If values fell below zero, they were recoded to zero. Only four such violation at the upper end of the nitrate distribution (at 22.9 mg/l, 23.6 mg/l, 23.7 mg/l and 24.2 mg/l) were observed for the reduced risk question. These violations are attributed to the restrictive nature of using polynomial functions for the analysis.

[11] Values depicted in Figure 5 were obtained by interpolating the WTP for 25 percent changes assuming local linearity in damages. To avoid extrapolation beyond the original change in nitrate levels, the analysis is truncated at 4 mg/l.

example, a shift from 20 to 15 mg/l would still leave them in a definitely unsafe zone. In spite of these majority feelings however, a small positive reduction is observed over the range because some individuals still perceive the reduction to improve their health probabilities.

SUMMARY AND IMPLICATIONS

This paper argues that if risk perceptions are a subjective function of reference exposure levels, then damage functions should be conditional upon the exposure level. In a case study of nitrates, subjective safety perceptions of nitrate exposure levels were found to be related to baseline exposure levels. Individuals in different baseline exposure categories had different exposure-safety transformation functions on average. In spite of their differences, however, the exposure-safety transformation for all reference groups responded to government standards as predicted by a threshold model.

Adopting the conditional damages perspective, functions of incremental damages and benefits were estimated and found to have areas of convexity and concavity. Importantly, perceived benefits and damages appear to be based on the information provided, and the benefits were highest for exposure levels for which incremental reductions in exposure approach government health standards. A second result from the empirical analysis is that incremental damages and benefits diminish after reaching a peak at an intermediate nitrate level. This result contrasts with the conventional approach to damage assessment which suggests that WTP for a small reduction in exposure monotonically increases with exposure levels. The implication of this finding is that the greatest benefits from intervention will occur at some intermediate level of exposure.

The eventual decline of incremental WTP does not appear to be attributed to averting actions that have been undertaken. Only a small portion of the households in the study had undertaken averting activities and the coefficients on these actions were insignificant. Instead, the upper bounds on WTP are attributed here to the opportunity for substitution through the establishment of a choke price and the fact that the incremental shifts in exposure levels do not greatly affect perceived safety at high levels of exposure.

REFERENCES

Baird, B. N. R., 1986. "Tolerance for Environmental Health Risks: The Influence of Knowledge, Benefits, Voluntariness, and Environmental Attitudes", *Risk Analysis*, 6: 425-435.

Cameron, T. A., 1988. "A New Paradigm for Valuing Non-Market Goods Using Referendum Data: Maximum Likelihood Estimation by Censured Logistic Regression", *Journal of Environmental Economics and Management*, 15:355-379.

Conrad, J. M. and L. J. Olson, 1992. "The Economics of a Stock Pollutant: Aldicarb on Long Island", *Environmental and Resource Economics*, 2:245-258.

Crocker, T. D., B. A. Forster, and J. F. Shogren, 1991. "Valuing Potential Groundwater Protection Benefits", *Water Resources Research*, 27:1-6.

Edwards, S. F., 1988. "Option Prices for Groundwater Protection," *Journal of Environmental Economics and Management*, 15:475-487.

Hamilton, L. C., 1985. "Concern About Toxic Wastes: Three Demographic Predictors", *Sociological Perspectives*, 28:463-486.

Kahneman, D. and A. Tversky, 1979. "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica*, 47: 263-291.

Kahneman, D. and J. L. Knetsch, 1992. "Valuing Public Goods: The Purchase of Moral Satisfaction", *Journal of Environmental Economics and Management*, 22:57-70.

Kask, S. B. and S. A. Maani, 1992. "Uncertainty, Information, and Hedonic Pricing", *Land Economics*, 68:170-184.

Kopp, R. J., and V. K. Smith, eds., 1993. *Valuing Natural Assets: The Economics of Natural Resource Damage Assessment*, Resources for the Future, Washington D.C.

Jones-Lee, M., 1974. "The Value of Changes in the Probability of Death or Injury" *Journal of Political Economy*, 82: 835-849.

Jones-Lee, M. W., M. Hammerton, and P. R. Philips, 1985. "The Value of Safety: Results of a National Sample Survey", *The Economic Journal*, 95:49-72.

Lichtenstein, S., P. Slovic, B. Fischhoff, M. Layman, and B. Combs, 1978. "Judged Frequency of Lethal Events", *Journal of Experimental Psychology: Human Learning and Memory*, 4: 551-578.

Loomis, J. B. and P. H. duVair, 1993. "Evaluating the Effect of Alternative Risk Communication Devices on Willingness to Pay: Results from a Dichotomous Choice Contingent Valuation Experiment", *Land Economics*, 69:287-298.

McClelland, G. H. *et al.*, 1992. *Methods for Measuring Non-Use Values: A Contingent Valuation Study of Groundwater Cleanup*, Center for Economic Analysis, University of Colorado, Boulder.

Mitchell, R. C., and R. T. Carson, 1989b. *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Resources for the Future, Washington D.C.

Page, T., and J. Ferejohn, 1974. "Externalities as Commodities", *American Economic Review*, 64:454-459.

Poe, G. L., 1993. *Information, Risk Perceptions and Contingent Values: The Case of Nitrates in Groundwater*, Doctoral Dissertation, University of Wisconsin-Madison.

Portage County Groundwater Management Plan, Volume 1, 1987.

Quiggen, J., 1992. "Risk, Self Protection and Ex Ante Economic Value-Some Positive Results", *Journal of Environmental Economics and Management*, 23:40-53.

Repetto, R., 1981. "The Economics of Visibility Protection: On a Clear Day You Can See a Policy", *Natural Resources Journal*, 21:355-70.

Repetto, R., 1987. "The Policy Implications of Non-Convex Environmental Damages: a Smog Control Study", *Journal of Environmental Economics and Management*, 14: 13-29.

Römer, A. U., and W. W. Pommerehne, 1990. *Valuing Reductions of Public Risks: The Case of Hazardous Waste*, Paper prepared for the Fifth Karlsruhe Seminar on Models and Measurement of Welfare and Inequality, August 12-19.

Schultz, S. D. and B. E. Lindsay, 1990. "The Willingness to Pay for Groundwater Protection", *Water Resources Research*, 26:1869-1875.

Shitaba, H. and J. S. Winrich, "Control of Pollution when the Offended Defend Themselves", *Economica*, 50:425-37.

Shogren, J. F., and T. C. Crocker, 1991, "Risk, Self-Protection and Ex Ante Economic Value", *Journal of Environmental Economics and Management*, 20:1-15.

Slovic, P., 1987. "Perception of Risk", *Science*, 236:280-285.

Slovic, P., B. Fischhoff, and S. Lichtenstein, 1979. "Rating the Risks", *Environment*, 21:14-39.

Smith, V. K., 1986. "Nonuse Values in Benefit Cost Analysis", *Southern Economic Journal*, 54:19-26.

Smith, V. K., and W. H. Desvousges, 1985. "Averting Behavior: Does It Exist?", *Economics Letters*, 20:291-296.

Smith, V. K., and W. H. Desvousges, 1987. "An Empirical Analysis of the Economic Value of Risk Changes", *Journal of Political Economy*, 95:89-114.

Smith, V. K., W. H. Desvousges, F. R. Johnson, A. Fisher, 1990. "Can Public Information Programs Affect Risk Perceptions?", *Journal of Policy Analysis and Management*, 9:41-59.

Smith, V. K., and F. R. Johnson, 1988. "How Do Risk Perceptions Respond To Information? The Case Of Radon", *The Review Of Economics and Statistics*, 70(1):1-9.

Starr, C., 1969. "Social Benefit versus Technological Risk", *Science*, 165:1232-1238.

Sun, H., J C. Bergstrom, and J. H. Dorfman, 1992. "Estimating Benefits of Groundwater Protection Control", *Southern Journal of Agricultural Economics*, 63-71.

Van Liere, K. D. and R. E. Dunlap, 1980. "The Social Bases of Environmental Concern: A Review of Hypotheses, Explanations and Empirical Evidence", *Public Opinion Quarterly*, 44:181-197.

Viscusi, W. K., 1989. "Prospective Reference Theory: Towards an Explanation of the Paradoxes", *Journal of Risk and Uncertainty*, 2:235-264.

Viscusi, W. K., W. A. Magat, and J. Huber, 1987. "An Investigation of the Rationality of Consumer Valuations of Multiple Health Risks", *Rand Journal of Economics*, 465-479.

Vlek, C. and P. J. Stallen, 1981. "Judging Risks and Benefits in the Small and in the Large", *Organizational Behavior and Human Performance*, 28:235-271.

Weinstein, M. C., D. S. Shepard, and J. Plisken, 1980. "The Economic Value of Changing Mortality Probabilities: A Decision Theoretic Approach", *Quarterly Journal of Economics*, 94: 373-395.

Xepapadeas, A. P., 1992. "Environmental Policy Design and Dynamic Nonpoint Pollution", *Journal of Environmental Economics and Management*, 23:22-40.

Zeckhauser, R. and A. Fisher, 1976. *Averting Behavior and External Diseconomies*, Discussion Paper, Kennedy School of Government, Havard University.

Table 1: Reference Levels and WTP for Hazardous Waste Risk Reduction from Römer and Pommerehne

"Objective" Risk Change	Intitial "Objective" Risk		t Value[a,b]
	0.0005	0.0001	
0.0005 --> 0.0001	37.06	NA	NA
0.0001 --->0.00005	10.76	28.25	4.35***
0.00005 ---> 0.000025	NA	10.29	NA

[a] one-sided test   [b] *** = 1% level of significance

Table 2: Descriptive Statistics of Model Variables

Variable	Description	Mean [n=205]	Mean [n=208]
LIVEPAST	Categorical variable for number of years of residence in Portage County: 0 = less than 1 year; 1 = 1 to 5 years; 2 = 6 to 10 years; 3 = 11 to 15 years; 4 = over 15 years.	2.35 (1.05)	2.34 (1.05)
OWNAGE	Categorical variable: 1 = less than 18; 2 = 18 to 44; 3 = 45 to 64; 4 = 65 or older.	2.71 (0.77)	2.70 (0.77)
DSEX	Binary variable for sex of respondent: 0 = male; 1 = female.	0.39 (0.49)	0.38 (0.49)
DCOLLEGE GRAD	Binary variable for college graduate: 0 = no; 1 = yes.	0.25 (0.44)	0.26 (0.44)
DFARM	Binary variable for involvement in farming: 0 = no; 1 = yes.	0.20 (0.40)	0.20 (0.40)
DAVTPERM	Binary variable for permanent averting activities of installing a purification system or carrying water from another source: 0 = no; 1 = yes.	0.04 (0.21)	0.04 (0.19)
DBOTWAT	Binary variable for purchase of bottled water for health reasons: 0 = no; 1 = yes.	0.03 (0.18)	0.03 (0.18)
NON-USE	Categorical variable or nitrate health concerns about other people living today and future generations: Ranging from 2 = not concerned to 8 = extremely concerned.	6.70 (1.31)	6.71 (1.31)
NITRATE	Nitrate level	5.93 (4.99)	5.98 (4.99)
$NITRATE^2$	Squared nitrate level	59.95 (102.89)	60.58 (102.46)
$NITRATE^3$	Cubed nitrate level	839.73 (2132.35)	844.46 (2119.10)

Table 3a: Polynomial Functions of Incremental Benefits (25% reduction)

	Linear Nitrates Only	Full Linear	Quadratic Nitrates Only	Full Quadratic	Cubic Nitrates Only	Full Cubic
CONSTANT	61.64 (50.78)	-72.11 (217.34)	-35.02 (77.63)	-154.38 (224.51)	29.35 (48.19)	-115.10 (220.31)
LIVEPAST		-41.91 (31.50)		-41.58 (31.46)		-44.85 (32.17)
OWNAGE		-116.96 (46.21)**		-122.50 (46.99)***		-121.58 (47.91)**
DSEX		-87.24 (65.43)		-73.19 (65.74)		-86.08 (66.91)
DCOLLEGE GRAD		54.83 (69.34)		40.83 (69.87)		34.51 (71.49)
DFARM		-132.15 (77.95)*		-150.23 (79.59)*		-153.32 (81.04)*
DAVTPERM		254.77 (225.44)		240.28 (216.03)		197.35 (234.66)
DBOTWAT		-42.89 (161.05)		-52.81 (162.53)		-78.68 (167.73)
NON-USE		95.22 (28.90)***		96.55 (28.95)***		100.45 (29.63)***
NITRATE	12.47 (6.49)*	7.84 (6.44)	46.16 (20.87)**	38.98 (20.78)*		
$NITRATE^2$			-1.71 (0.99)*	-1.57 (0.99)	4.92 (1.67)***	4.26 (1.73)**
$NITRATE^3$					-0.22 (0.08)***	-0.20 (0.09)**
k	213.03 (41.28)***	173.75 (32.21)***	209.73 (40.12)***	173.00 (31.84)***	207.66 (39.84)***	175.16 (32.25)***
n	221	208	221	208	221	208
χ^2	47.58	82.81	50.84	85.57	55.34	89.12
McFadden R^2	0.16	0.29	0.17	0.30	0.18	0.31

Notes: Asymptotic Standard Errors in (). Significance levels are denoted * (10 percent), ** (5 percent) and *** (1 percent)

Table 3b: Polynomial Functions of Incremental Damages (avoid a 25% increase)

	Linear Nitrates Only	Full Linear	Quadratic Nitrates Only	Full Quadratic	Cubic Nitrates Only	Full Cubic
CONSTANT	24.38 (51.48)	-427.09 (217.28)	-34.94 (81.89)	-463.51 (226.51)	23.34 (49.98)	-415.81 (215.25)*
LIVEPAST		54.02 (32.77)		53.91 (32.76)		53.02 (32.48)
OWNAGE		-84.93 (41.37)**		-88.86 (42.12)**		-88.52 (41.70)**
DSEX		-124.06 (64.73)*		-118.17 (65.41)*		-128.90 (65.09)**
DCOLLEGE GRAD		259.75 (71.56)***		258.51 (71.84)***		258.46 (71.72)***
DFARM		-9.35 (70.68)		-16.04 (71.74)		-13.52 (71.18)
DAVTPERM		-14.35 (197.74)		-13.32 (199.13)		-32.73 (203.07)
DBOTWAT		424.96 (182.17)**		434.08 (185.32)**		415.89 (184.99)**
NON-USE		82.00 (26.18)***		83.30 (26.47)***		83.44 (26.32)***
NITRATE	15.49 (6.56)**	13.62 (6.72)**	35.23 (21.09)*	27.18 (20.13)		
$NITRATE^2$			-1.01 (1.01)	-0.73 (1.00)	3.58 (1.57)**	2.94 (1.49)**
$NITRATE^3$					-0.15 (0.08)*	-0.12 (0.07)*
k	197.00 (36.04)***	156.31 (28.08)***	201.36 (37.65)***	156.78 (28.38)***	200.98 (37.67)***	155.74 (28.33)***
n	218	205	218	205	218	205
χ^2	49.78	93.42	50.86	93.95	52.25	94.95
McFadden R^2	0.17	0.34	0.17	0.34	0.18	0.34

Notes: Asymptotic Standard Errors in (). Significance levels are denoted * (10 percent), ** (5 percent) and *** (1 percent)

Figure 1: Reference-Based Risk Perceptions [g(P|P)] and Subjective Expected Utility (EU) Loci

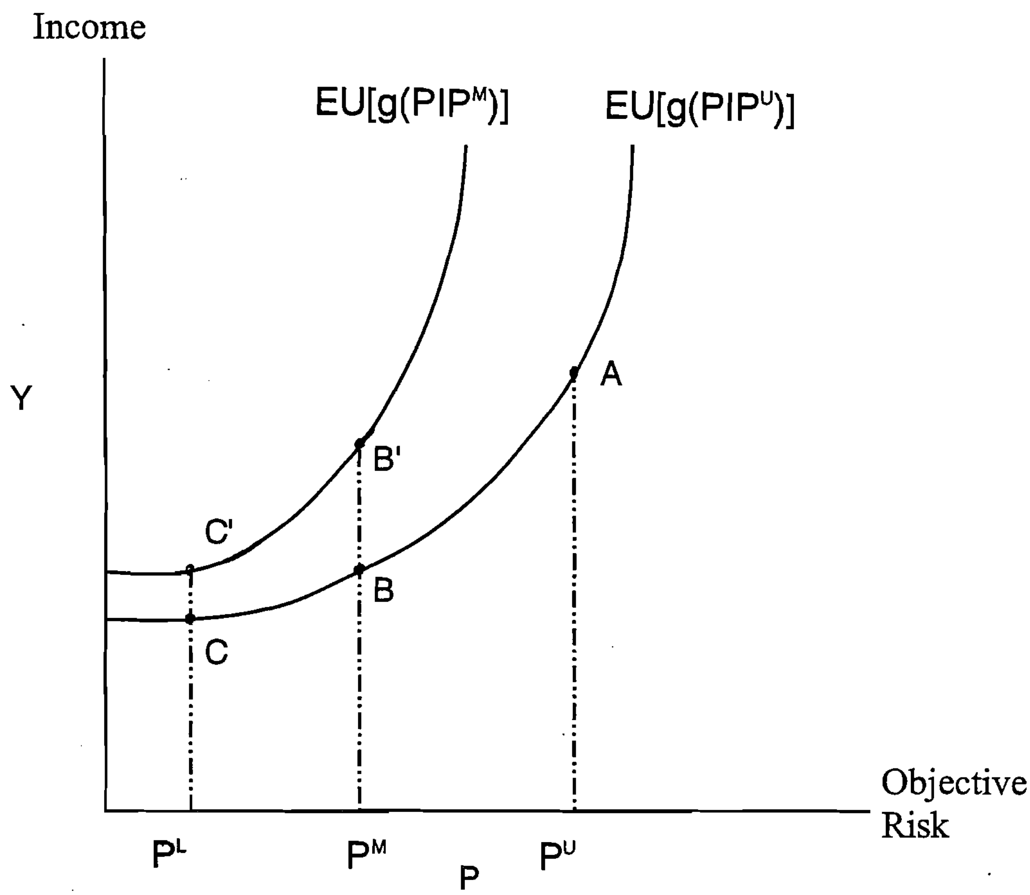
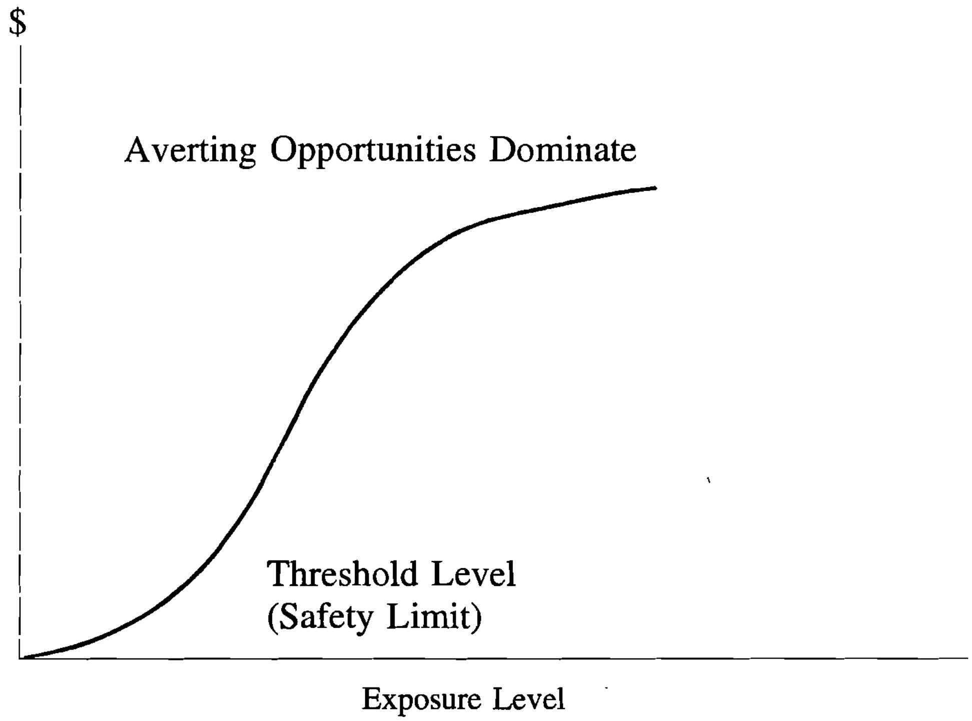


Figure 2: Plausible Damage Functions

Total



Marginal or Incremental

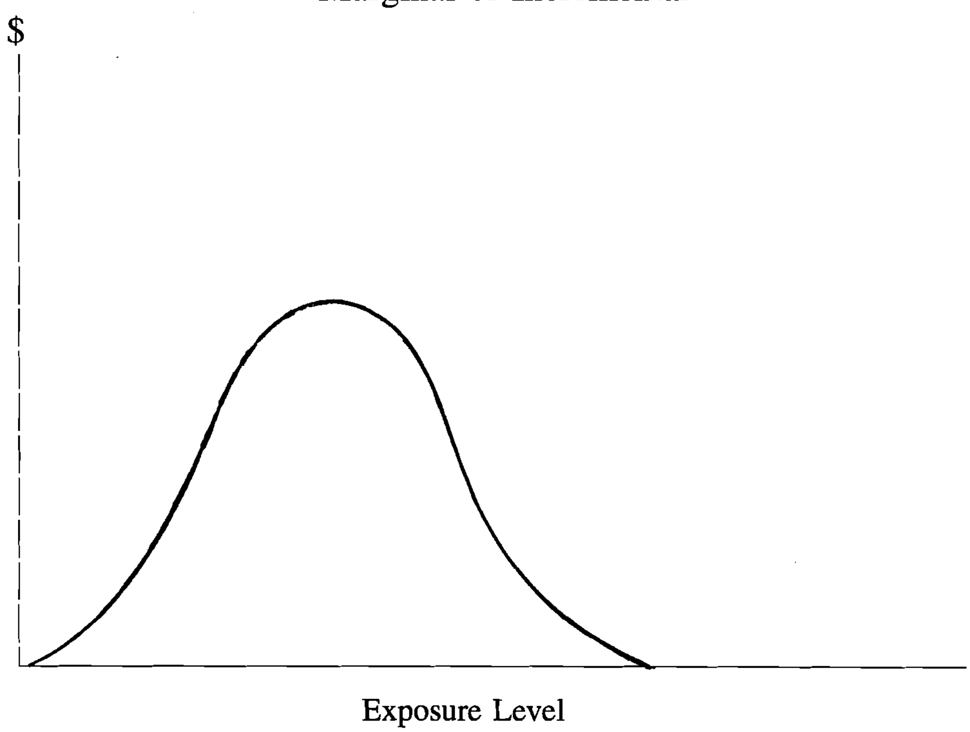
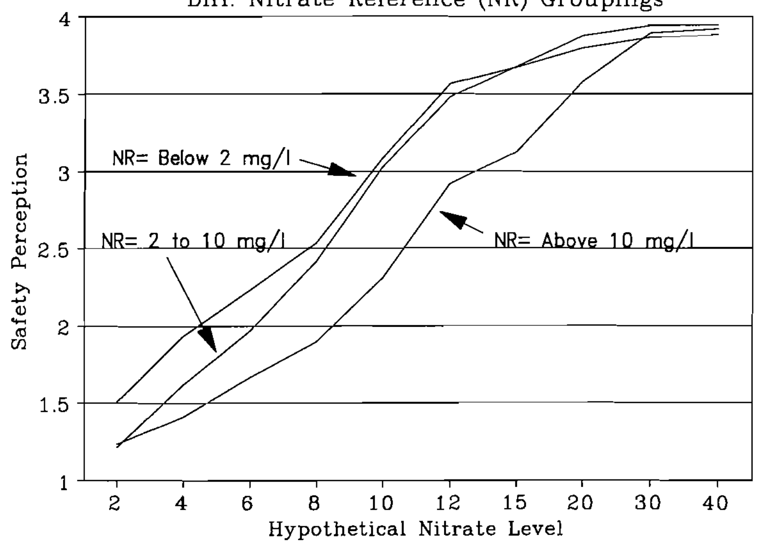


Figure 3: Safety Perceptions for
Diff. Nitrate Reference (NR) Groupings



Note: 1 = Definitely Safe, 2 = Probably Safe, 3 = Probably Not Safe, 4 = Definitely Not Safe

Figure 4: Incremental Damage and
Benefits Functions (25 Percent Change)

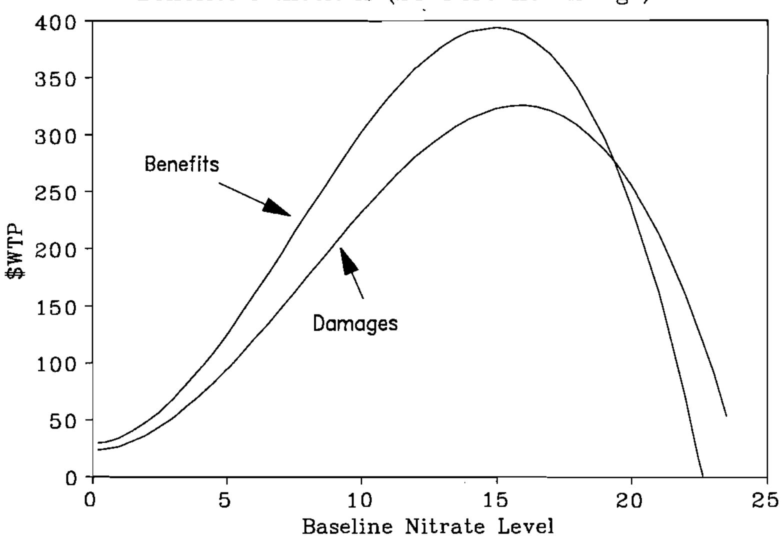
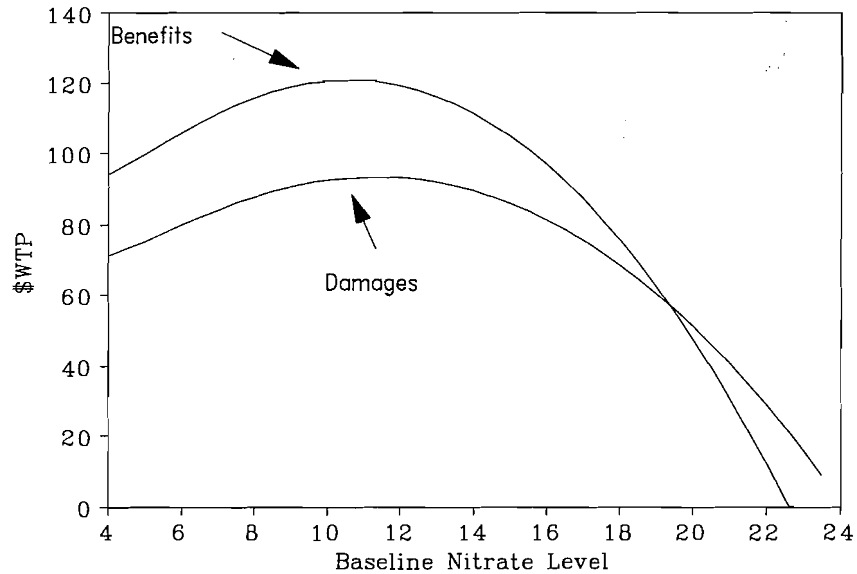


Figure 5: Marginal Damage and Benefit Functions (imputed)



MAKING SENSE OF SUSTAINABILITY

Alan Randall*

The Ohio State University

Contrary, perhaps, to the impressions of non-specialists, there already exists a substantial economic literature on sustainability. There is a considerable economic-theoretic literature and a considerable prescriptive literature. The intersection, even, is non-empty; that is, some of the theoretical literature takes seriously the task of prescription, and some of the prescriptive literature is sensitive to what can be learned from economic theory. It is true there is little empirical literature; but I find it hard to be critical about that: it is not easy to imagine what a meaningful economic-empirical literature about sustainability would look like.

The diagnostic and prescriptive literature appears at first glace noisy and discordant. Diagnoses range from simple market failures to modern lifestyles incompatible with the carrying capacity of the planet. Policy prescriptions run the gamut from correction of market failures to elimination of discounting, intergenerational reassignment of entitlements, optimal re-investment rules for natural resource rents, and a safe minimum standard of conservation; and that is just from relatively mainstream resource economists. Some of our ecological economist colleagues would extend the range of prescriptions to include "robust strategies" emphasizing resiliency, and radical restructuring of the modern consumer economy and society.

In trying to make some sense of all this disagreement about diagnosis and prescription, perhaps the place to start is with the theoretical literature.

---

*I have benefitted greatly from a dialogue with Mike Farmer that continues beyond the research for his dissertation (1993).

Economic Theory and Sustainability

What, exactly, are the theorists concerned about sustaining? The literature suggests at least five different sustainability goals.

Sustainability Goals

1. *Maintaining Welfare, or Aggregate Output.* A reasonable goal is to sustain welfare across the generations. The Bruntland Commission's definition--meet(ing) the needs of the present without compromising the ability of future generations to meet their own needs (WCED 1987)--would surely be satisfied by any arrangement that succeeds in maintaining welfare for the indefinite future.

Solow's famous (1974) formulation addresses aggregate output:

$$\frac{Y}{L} = e^{t}\left(\frac{D}{L}\right)^{g}\left(\frac{K}{1}\right)^{1-g-h}$$

where Y is aggregate output, L is labor (i.e., population, such that dividing by L puts things in per capita terms), D is natural resources, K is reproducible capital, technology is Cobb-Douglas, and t is the rate of technological progress. However, output is aggregated in such a way that maintaining Y/L is, in effect, maintaining welfare.

2. *Maintaining the Stock of Capital.* This goal which addresses D plus K, (i.e., societal wealth properly indexed and aggregated) arises from the Solow view of the world; especially, from his favorable assumptions about the substitutability of D and K. To meet this goal, some type of Hartwick (1977, 1978) rule is followed: the scarcity rents from natural resources exhaustion must be re-invested in reproducible capita. The purpose of such a rule is to maintain the productive capacity of society which, if accomplished, would maintain welfare.

Notice immediately that the Hartwick rule is either tautological or wrong. If D and K are excellent substitutes (e.g., as would be the case with CES aggregate production technology and substitution elasticity ≥ 1), if K and Y are aggregated and indexed according to optimal pricing rules, and if resource rents reflect correctly the value of incremented scarcity due to extraction, then that rule is correct by definition. Otherwise, satisfying the Hartwick rule is insufficient to sustain welfare.

3. *Maintaining Natural Resources.* If natural resources really are different, i.e., D and K are not very good substitutes, then sustainability policy has to be targeted at D itself. Daly (1990), and Pearce and Turner (1990) are among the economists who have tried to delineate policies addressed specifically to D. El Serafy (1989) has proposed a rule requiring that habitats and biotic resources not be used beyond their long-run regenerative capacity, and exhaustible resources be depleted no more rapidly than they can be replaced by sustainable harvest of renewables. Barbier, et al. (1990) propose a policy of compensatory projects, such that non-sustainable harvest of a particular resource is compensated by some particular D-enhancing project in order to sustain aggregate D.

While the standard growth model characterizes D as natural resources for production, it is well to remember the importance of nature for assimilating wastes. Some of the major sustainability issues currently on the public mind--e.g., global warming, and depletion of the ozone layer--concern the waste assimilation capacity of D. The maintenance of natural resources may require constraints on release of wastes.

4. *Ecological Sustainability.* If biotic resources really are importantly different from K and from, say, mineral deposits, then sustainability policy should be targeted toward biotic, or ecological sustainability (e.g. Common and Perrings 1992). Such an approach may well require radical re-thinking of how economists model sustainability issues; and it may well suggest radical restructuring of modern consumer society. Arguments to support these kinds of approaches are likely to involve not just the modeling assumptions but also the ethical stance of biocentrism or "deep ecology" (see e.g., Taylor 1981, 1983).

5. *Preservation of Particular Natural Resources.* Regardless of one's position concerning aggregate Y, D, and K, there may be particular natural phenomena--geological formations, habitats, ecological associations, or species--that one wants to see preserved for the future. Preservation arguments of this kind seldom hinge on urgent concerns about human survival (or, if they do, they logically collapse into one of the above four categories). Preservation motives range from the utilitarian (these things provide pleasure indirectly or directly), to claims of intrinsic value (they have a good of their own), to claims that they have rights that we are obligated to respect. One commonality, however, is the premise of uniqueness, i.e., that the thing to be preserved has little in the way of acceptable substitutes.

Since preserving certain particular natural resources is acceptable (although likely for different reasons) to proponents of the first four kinds of sustainability goals, I will first concentrate on goals one through four. I will, however, eventually return to issues concerning preservation of natural resources.

Modeling Assumptions

The choice of sustainability goals, and the modeling results concerning attainability of any particular goal, depend on modeling assumptions.

Cake or Corn: Is Production Modeled Explicitly? Cake-eating models deal with the optimal depletion of a given endowment, and generate only one robust result: a society that discounts the value of future consumption will choose a consumption path declining with time. Within one's own life, such a choice might be termed myopic. In a multi-generational context, such selfish behavior can be supported only by a positional dictatorship of the present generation (Ferejohn and Page 1977). From the perspective of sustainability, however, none of this is very interesting: a cake-eating universe is inherently unsustainable, and the kinds of discussions one can base upon such models have an unrelievedly pessimistic tone.

At the opposite end of the optimism-pessimism scale, Solow (1974) provides a model in which society could conceivably maintain its welfare across indefinitely many generations even though it uses exhaustible resources. Solow's model explicitly considers production, but to Solow (1974), production is greatly facilitated by Cobb-Douglas technology and perfectly-divisible D.

An intermediate position considers a natural resource that is capable of regeneration, within the bounds set by biological possibilities. Future prospects are influenced by the regenerative capacity of the natural resource, as well as the degree to which reproducible K can substitute for it.

Substitutability. In comparing the first four sustainability goals, perhaps the first thing that strikes one is the importance of assumptions about substitutability. Models addressed to the first two goals typically assume generous substitution between particular resources and between aggregate D and K. Maintaining welfare clearly permits a broad range of substitution in consumption, as does the concept of aggregate output. Generous substitutability is assumed in production, such that output can be maintained even as the composition of aggregate capital shifts markedly. While seldom modeled explicitly, it is clear from the discussion in this literature that technology is assumed to progress over time and to respond to relative scarcity so that its progress is tilted toward increasing the substitutability of plentiful resources for those that are scarce. In some treatments, K is clearly intended to include human capital and to embody progressing technology.

Analysts who are more impressed with the limits of substitutability, gravitate to sustainability objectives (3) and (4). They see the need to focus sustainability policy specifically on maintaining natural resources and/or biotic resources.

Substitutability can, of course, be a matter of more than tastes and technology. Some of the literature in environmental ethics and most of the "deep ecology" literature suggests ethical limits on substitution: to substitute the artificial for the natural and be just as happy may be, *ipso facto*, an indication of depravity.

Regeneration of Biotic Resources. It is common for economists to model regeneration of biotic resources as a function, often sigmoid in shape. I will do some of that, later in this paper. To conceptualize uncertainty, I assume that the regeneration function is not deterministic but can be represented as a confidence band. The more risk-averse among us can focus mainly on the lower boundary of that band.

While for economists that is a considerable concession to existential uncertainty, many ecologists believe that in reality much less is known about the regeneration of natural populations. While economists seek point solutions identified by familiar tangents to regeneration curves, ecologists are more likely to examine the resiliency of the populations and to seek robust policy solutions that perform reasonably well over a broad range of conditions.

Single-agent or Structural Models? Models in the Ramsey-Solow tradition are single-agent models. There is no division in roles, e.g. producer, consumer, government; and no populations of folk in different circumstances who might be motivated to trade, so that prices may emerge.

Recently, Howarth and Norgaard (1990) and Farmer (1993) have developed conceptual analyses in which the structure of succeeding generations is explicitly modeled. These models produce insights about resource prices, discount rates, and endogenous incentives for rationing and resource conservation that are unattainable with singe-agent models.

Lessons From an Over-lapping Generations Model

Farmer (1993) constructed an overlapping generations model along the following lines. At any time, there are three generations living (young, y; middle-aged, m; and retired, r) For any individual, an optimal life-plan maximizes

$$U(C_y) + U(C_m) + U(C_r)$$

(where C is aggregate consumption), subject to production technology, the regeneration function for D, and various accounting restrictions: the young borrow K and buy D; the middle-aged lend K and sell D; the retired just consume; production combines D and K to produce (more) K; all consumption is taken from K; all budgets balance; and materials balance.

The model starts with initial endowments of D and K, and determines resource allocation, consumption, and prices endogenously, as the generations trade with each other and succeed each other. In the model, all agents have perfect foresight. This is not stacking the deck: much of the previous literature worries that selfish agents, even with perfect foresight, may choose an unsustainable consumption path. Farmer's agents are selfish, rather than altruistic; intergenerational altruism is much to be encouraged and can only help in the quest for sustainability, but it would be stacking the deck to assume it.

This model enables us to critique four rather standard prescriptions for sustainability.

Discounting Is Not the Problem, and Discount Rate Repression Is Not the Solution. It is perhaps the most enduring of myths that a society which discounts future production and costs *ipso facto* sacrifices future welfare, and therefore violates reasonable requirements for intergenerational equity (Young 1992). Note that the individuals in Farmer's model maximize welfare summed, undiscounted, across the three life-stages. The individuals are neutral with respect to time preferences about consumption. Nevertheless, positive interest rates emerge endogenously. Why? Because capital is scarce and productive, and the young have to buy (borrow) it.

In Farmer's model, future prospects depend on what is assumed about initial endowments, the substitutability of D and K, and the regeneration of D. A considerable range of outcomes is possible: welfare may be increasing or decreasing over time; resource crises may occur, even with perfect foresight. In cases where future prospects are for declining welfare, it may be tempting to blame the positive interest rates that emerge endogenously, and to prescribe discount rate repression in order to raise future consumption. But that would be the wrong diagnosis and the wrong prescription: regardless of whether the consumption path is increasing or decreasing, a policy of interest rate repression would only make things worse for the future. Furthermore, this result has nothing to do with any positional dictatorship of the present generations. Unborn future generations would prefer that those living now face incentives to save, and to select only those investments that pass a net present value test.

Entitling Future Generations Will Help Them Less Than One Might Think. Recently, Bromley (1989) proposed that the problem of sustainability could be solved by an appealingly simple yet effective instrument: a reassignment of property rights to future generations. This approach would be effective: a future generation protected by property rights would have veto power over earlier-generation actions that might threaten its welfare. It would be simple: the property rights reassignment to the future would be once-and-for-all (although it would require a momentous public decision to actually make such a change); and enforcement of the reassigned property rights would proceed routinely, as does current enforcement of currently-assigned property rights.

Howarth and Norgaard (1990) endorse this proposal, based on their analysis with a two-generation, overlapping generations model, in which prices are given exogenously. They start by examining trade between adjacent generations, given that property rights are first reassigned from the older to the younger. Then, by induction, they consider entitlement of distant future generations.

Both Bromley (1989) and Howarth and Norgaard (1990) are alert to the Coase theorem, which would suggest that reassignment of property rights (even across generations) would have less impact on resource allocation than one might think. Nevertheless, they conclude that Coasian concerns do not undermine the validity of their proposal.

In Farmer's model, intergenerational trading opportunities are much more complete than in the Howarth-Norgaard model. With three generations, asset and capital markets are completely characterized, and prices are endogenized. Production responds to prices, and prices respond to demands. The Coase theorem, properly interpreted, says something like: the fewer are the impediments to trade, the more nearly are resource allocation outcomes insensitive to the initial assignment of rights. Farmer's results conform to the Coasian insight. The assignment of property rights to each successive young generation at birth provides only modest protection for the immediate unborn generation; the effect on more distant generations is indeterminant. In cases where the model predicts that current consumption levels are unsustainable, the reassignment of property rights is typically insufficient to reverse that outcome. To express it more formally, the Howarth-Norgaard finding -- that reassignment of property rights to future generations is sufficient to secure future welfare -- is not attainable as a general equilibrium result.

Hartwick Rules Are Not Policy Prescriptions. Hartwick rules require that Hotelling (i.e., scarcity) rents from exhaustible resource extraction be re-invested. I have argued, above, that the claim that Hartwick rules assure sustainability is either tautological or wrong. Here, I address their serviceability in prescription.

There's no assurance that a Solow single-agent economy will generate the prices that validate the Hartwick tautology (Krautkraemer et. al. 1994). There are enormous obstacles to, first, measuring the rents from resource depletion and, then, overcoming the incentive problems in controlling capital investment to ensure that the ex ante and ex post value of national wealth is unchanged. Further, the problem of price formation, in the structural sense, is ignored. To borrow an example from Mike Farmer, Hartwick rules assume we can chop down an entire rainforest and reinvest the rents in some reproducible asset of equal value, all without affecting the prices of either asset. It is a policy without an implementation prescription.

Safe Minimum Standard Policies Have Some Promise. Randall (1991) and Randall and Farmer (1994) have argued that a policy rule to allocate natural resources on the basis of efficiency criteria, but always subject to a safe minimum standard (SMS) of conservation (Ciriacy-Wantrup 1968) would be taken seriously by ethicists operating from a broad range of philosophical perspectives. The SMS is a constraint adopted for good reason, and the constraint itself can be abandoned if the cost of enforcing it becomes intolerably high (Bishop 1978). Here, I plan to address three related issues: the role of a SMS constraint in policy for sustainability, principles for setting the SMS, and the problem of implementation.

To address these questions, consider a simple two-period diagram. Assume D is renewable, that is, D withheld from production in one period regenerates by the next period. If S_t is the stock of D withheld from production in period t, the regeneration function traces the relationship between S_t and S^a_{t+1}, the amount of D available in the next period. In a two-period diagram, the line of slope=1 starting from the origin is diagnostic: at points above that line, S^a_{t+1} exceeds S_t so that the natural resource is at least potentially sustainable; but at points below the line, the natural resource will eventually be exhausted even if none of it is used in production (Figure).

Assume perfect foresight and efficient markets in Y, D, and K. An interesting question is whether natural resource "crises" (i.e., situations where scarcity of natural resources threatens the sustainability of adequate consumption levels for the human population) are possible. Assume that D and K are not perfect substitutes and that factor-specialization is penalized in production.

If the regeneration function is always concave and lies above the line of slope = 1 for a range of values of S_t, it will have a steep positive slope near the origin. In this case, the market economy provides very strong defenses against resource crises: the price of D will grow very large as the resource nears exhaustion, and any S_t conserved as a result of this incentive will regenerate generously $(S^a_{t+1} > S_t)$.

The sustainability question becomes more interesting if the natural resource regeneration function is sigmoid (Figure). If less than S_{min} is withheld from production in each period, natural resource exhaustion is inevitable. The optimal stock to carry forward is S^*_t, at which point the steady-state efficiency condition, 1+r = 1+h, holds (where r is the marginal efficiency of capital, and h is the marginal regeneration rate of the natural resource) and D^*_t may be used in production in each period.

Interpreting S_{min} as the minimum standard (i.e., the minimum carry-over stock to assure resource regeneration), the idea of a *safe* minimum standard invokes uncertainty. Assume that the regeneration function is stochastic and that its lower bound is traced by the dashed curve (Figure). Then, if $\widetilde{SMS}$ is withheld from production in each period, resource exhaustion will be avoided, even in the worst case with respect to resource regeneration. We take $\tilde{SMS}$ as what is meant by the term safe minimum standard in the literature; we would call it safe minimum standard of *preservation.*

$\tilde{SMS}$ sustains the resource (and that may satisfy some preservationists). But we have cast the issue as one of sustaining adequate consumption levels for the human population. Assume that D_{min} is the minimum allocation of natural resources to production that is required to sustain adequate consumption. Let each time period, *t*, represent a generation of people. Then, any generation that uses less than D_{min} suffers extreme deprivation (however that is defined). We identify $\hat{SMS}$ (Figure) as the minimum stock withheld from production that will provide D_{min} for each succeeding generation. Draws of D_{min} and regeneration of the stock are guaranteed. $\hat{SMS}$ is the safe minimum standard of *conservation.* While conservation of $\hat{SMS}$ is *required* to *assure* sustainability, the odds of doing better than that are working in favor of a society that abides by an $\hat{SMS}$ constraint: if regeneration turns out to be better than lower-bound, as it probably will, subsequent generations will be able to use more than D_{min} and/or conserve more than $\hat{SMS}$.

Let us pause at this point, to observe that some progress has been made in addressing the first two issues. Why might a SMS constraint be needed? The story that emerges from Farmer's model is generally favorable to the prospects of sustainability given fully functioning intergenerational markets. Nevertheless, there are no general-form guarantees. If initial endowments at too low, D-K substitutability and the regeneration of D are ungenerous, and/or the system is subject to uncertainty and experiences a run of bad luck, sustainability may be jeopardized. With sigmoid regeneration and required minimum draws of D, the system could find itself on a slippery slope. Some kind of SMS constraint could be invoked, in order to protect society against such outcomes.

How should the SMS be set? Randall and Farmer (1994) argue that the safe minimum standard should be set at $\hat{SMS}$, a more conservative level than one might expect. $\hat{SMS}$ allows for continuing harvest of D_{min}, to meet the minimal consumption requirements of present generations.

The remaining question concerns implementation. At the outset, observe that all pro-active sustainability policies raise implementation issues: I have not addressed implementation of discount-rate repression, entitling the future, or Hartwick rules, only because I have dismissed these policies for other reasons. An $\hat{SMS}$ rule requiring present society to conserve resources to avoid exhaustion in some (perhaps distant) future generation is not a sustainable equilibrium outcome; in other words there is no Lockean contract that would bind present society to abide by $\hat{SMS}$ for the benefit of distant future societies. Rather, $\hat{SMS}$ is a commitment that a society might undertake for ethical reasons.

D_{min} is defined as the natural resource draw necessary to avoid extreme deprivation for the current human society. One would expect a generation that inherited a natural resource stock less than $\hat{SMS}$ to nevertheless use at least D_{min}, risking resource exhaustion for some subsequent generation. To do otherwise would be to voluntarily accept self-sacrifice (to drink from the poisoned cup, as it were) for the benefit of future societies. In practical terms, that seems too much to ask.

Ethical theories offer only limited help, here. While many ethical systems would require individual self-sacrifice for the sake of principle or for the good of others, there seems little basis in ethical theory for obliging a society to sacrifice itself for the good of future societies.

An implementable safe minimum standard policy must seek to conserve not $\tilde{SMS}$ but $\hat{SMS}$. That is, it must seek to avoid placing any present or future society in a position where it must choose between sacrificing itself and dooming subsequent societies. In practical terms, a $\hat{SMS}$ policy would emphasize early warning, and early implementation of conservation policies that require only modest sacrifice on the part of each society. Since unilateral withdrawal from any intertemporal contract or obligation is always a possibility, conservationists have a strong interest in keeping the costs of conservation tolerably low. In addition, as Barbier and Markandya (1990) have suggested, some societies may have already passed the point of no return: sustainability could not be achieved with internal resources regardless of willingness to sacrifice for the future. It may be possible, however, for more asset-rich societies to subsidize these "basket cases" back to a sustainable path.

Practical Policies to Remote Sustainability

To this point, I have been concerned mostly to provide some guidance to the economic-theoretic considerations that help rationalize and systematize a sometimes discordant literature; to debunk some popular panaceas; and to explore the potential of policies incorporating a safe minimum standard of conservation. Now, I offer some commentary on practical policies to promote sustainability.

Population and Technology. Population and technology, and what might be expected concerning their growth, figure prominently in most discussions of sustainability. I have not ignored these issues, but one might need to look hard in order to find where I have treated them. Population was acknowledged, but then submerged immediately when I presented the Solow (1974) model in per capita terms. Solow's (1974) observation--that output per capita could be maintained so long as technological progress kept pace with population growth--serves merely to state the problem. Policies to control the growth of the human population and to encourage continuing technological progress are essential to any meaningful sustainability policy.

The analysis of a safe minimum standard of conservation made much of D_{min}, the minimum natural resource draw to protect present generations from deprivation, and rightly so. Nevertheless, the magnitude of D_{min}, is itself an issue of technology: D_{min} would be reduced by a technology that increased the substitutability of K for D. If the resource crisis concerned not D, generic natural resources, but particular natural resources, the range of possible substitutions is expanded to include other, less scarce, natural resources.

Mainstream economists are fairly optimistic that market forces tend to encourage technological progress and direct it toward increasing the substitutability of more available resources for those that are increasingly scarce. Nevertheless, a pro-active technology policy would provide some additional insurance.

Accounting for Resource Depletion. I have argued that intertemporal/intergenerational markets are more complete and more effective in assuring sustainability than is widely suspected. Furthermore

(I have argued), Hartwick rules--invest rents from natural resource depletion in reproducible capital assets--have problems with respect to theoretical coherence and implementability.

Nevertheless, the general idea of systematic accounting for natural resource depletion has much to recommend it. National accounts do not substitute for the incentives that actually allocate resources, but they may serve to motivate the political will essential for redirecting incentives. Natural-resource-exporting countries, such as Australia and New Zealand, are naturally torn between consuming and investing the proceeds from resource extraction; and exhortations to invest more and consume less cannot hurt.

Getting the Prices Right. Whatever optimism we gain from economic-theoretic models of sustainability must be sobered by the realization that such models assume that the standard market failures are (already) resolved, the prices are (already) right, and government stands ready always to implement public policy proposals that pass a benefit-cost test.

Policies for sustaintability *must build upon* the common sense recommendations of resource economics:

• Correct market failures, by implementing efficient institutions (see, e.g., Johnson 1992). Many of the most egregiously unsustainable policies and practices would fail tests for efficiency, as well as sustainability. Many of the most obvious market failures concern the generation and release of wastes that threaten sustainability as surely as does resource scarcity.

• Provide those conservation policies that pass a standard modern benefit cost test, i.e., one that measures willingness to pay for preference satisfaction without undue regard to observable market prices. Remember, the Randall-Farmer argument for an SMS rule addresses such a rule imposed as a constraint upon (*not substituted for)* policies that pass an efficiency test.

Getting Ahead of the Game. Our development of SMS concepts leads to a clear policy recommendation. Get ahead of the game. Implement conservation policy while it is still cheap, i.e., before the crises are upon us, before the train wrecks are imminent, while the sacrifices inherent in a serious conservation policy are still modest. That way, we can be averse to environmental risk, without paying an excessive price for our risk aversion. Furthermore, given that moral arguments can at best persuade others to adopt obligations, it is best that the obligations upon succeeding generations to conserve for the benefit of more distant generations involve only limited sacrifice.

This recommendation springs logically from our development of the case for a safe minimum standard of conservation. While our arguments for the SMS deviate only modestly from the path of mainstream economics, I believe the policy conclusion is fairly consistent with the "robust strategies" concepts that are emerging from ecology and ecological economics.

Preservation of Particular Natural Resources. Optimists and pessimists with respect to future welfare, capital accumulation and/or conservation of generic natural resources (D) agree that there are some particular natural resources that should be preserved, even as they may disagree as to exactly which ones fall into this category (Solow 1992). It seems that I have spent most of the last 25 years worrying about this problem. Not surprisingly, I could discuss this question in more detail than most audiences could bear. Mercifully, I will leave you with just one observation.

Development, it has often been observed, is the process of converting particular natural resources into reproducible capital. It is natural and healthy to worry about the risk that we might stumble into giving up too much that is rare and irreplaceable to gain that which is generic and reproducible. Arrayed against that risk is an opposite risk: we might reduce present and future welfare by restraining excessively the process we call "development". While this dilemma often seems insoluble, a strong economy not only allows the luxury of preserving environmental particulars, but also generates increasing demands for such preservation. It is easier when we afford it and when the citizenry is demanding that we do it. If the optimists are right, and welfare follows an increasing path, the demands for preserving particular natural treasures will only increase.

References

Barbier, E. B. and A. Markandya. 1990. "The Conditions for Achieving Environmentally Sustainable Development," *European Economic Review* 34: 659-669.

Barbier, E. B., A. Markandya, and D. W. Pearce. 1990. "Environmental Sustainability and Cost-Benefit Analysis," *Environment and Planning* 22: 1259-1266.

Bishop, R. C. 1978. "Economics of Endangered Species." *American Journal of Agricultural Economics* 60: 10-18.

Bromley, Daniel W. 1989. "Entitlements, Missing Markets and Environmental Uncertainty," *Journal of Environmental Economics and Management* 17: 181-194.

Ciriacy-Wantrup, S. von. 1968. *Resource Conservation: Economics and Policies,* (3rd ed.) University of California Division of Agricultural Sciences, Berkeley.

Common, M.and C. Perrings, 192. "Towards an Ecological Economics of Sustainability," *Ecological Economics* 6: 7-34.

Daly, H. 1990. "Some Operationed Principles of Sustainable Development," *Ecological Economics* 2: 1-6.

El Serafy, S. 1989. "The Proper Calculation of Income from Depletable Natural Resources," in Y. J. Ahmad, S. el Serafy, and E. Lutz, eds., *Environmental Accounting for Sustainable Development* (Washington, D.C., The World Bank).

Farmer, M. C. 1993. "Can Markets Provide for Future Generations?" Ph.D. dissertation, The Ohio State University, Columbus.

Ferejohn, and Talbot R. Page. 1978. "On the Foundations of Intertemporal Choice," *American Journal of Agricultural Economics* 60: 269-275

Hartwick, J.M. 1977. "Intergenerational Equity and the Investing of Rents from Exhaustible Resources," *American Economic Review* 67: 972-974.

Hartwick, J. M. 1978. "Exploitation of Many Deposits of an Exhaustible Resource," *Econometrica* 46: 201-217.

Howarth, Richard B., and Richard B. Norgaard 1991. "Intergenerational Resource Rights," *Land Economics* 66: 1-11.

Johnson, R. W. M., 1992. "Resource Management, Sustainability, and Property Rights in New Zealand" *Australian Journal of Agricultural Economics* 36: 167-186.

Krautkraemer, J., J. Pezzey, and M. A. Toman. 1994. "Economic Theory and Sustainability," in D. W. Browley (ed.) *Handbook of Environmental Economics*, Blackwell (in press).

Pearce, D. W., and K. Turner. 1990. *Economics of Natural Resources and the Environment*. London, Harvester Wheatsheaf.

Randall, A. 1991. "The Economic Value of Biodiversity." *Ambio: A Journal of the Human Environment* 20(2): 64-68.

Randall, A., and M. C. Farmer. 1994. "Benefits, Costs, and a Safe Minimum Standard of Conservation," in D. W. Bromley (ed.) *Handbook of Environmental Economics*. Blackwell (in press).

Solow, Robert M. 1974. "Intergenerational Equity and Exhaustible Resources," *Review of Economic Studies: Symposium on the Economics of Exhaustible Resources* 41: 29-45.

Solow, R. 1992. "An Almost Practical Step Toward Sustainability," Washington, *Resources for the Future, Inc.*

Taylor, P. 1981. "The Ethics of Respect for Nature," *Environmental Ethics* 3: 197-218.

Taylor, P. 1983. "In Defense of Biocentrism," *Environmental Ethics* 5: 237-243.

World Commission on Environment and Development (WCED). 1987, "Our Common Future," New York, *Oxford University Press.*

Young, R., 1992. "Evaluating Long-Lived Projects: The Issue of Intergenerational Equity," *Australian Journal of Agricultural Economics* 36: 207-232.

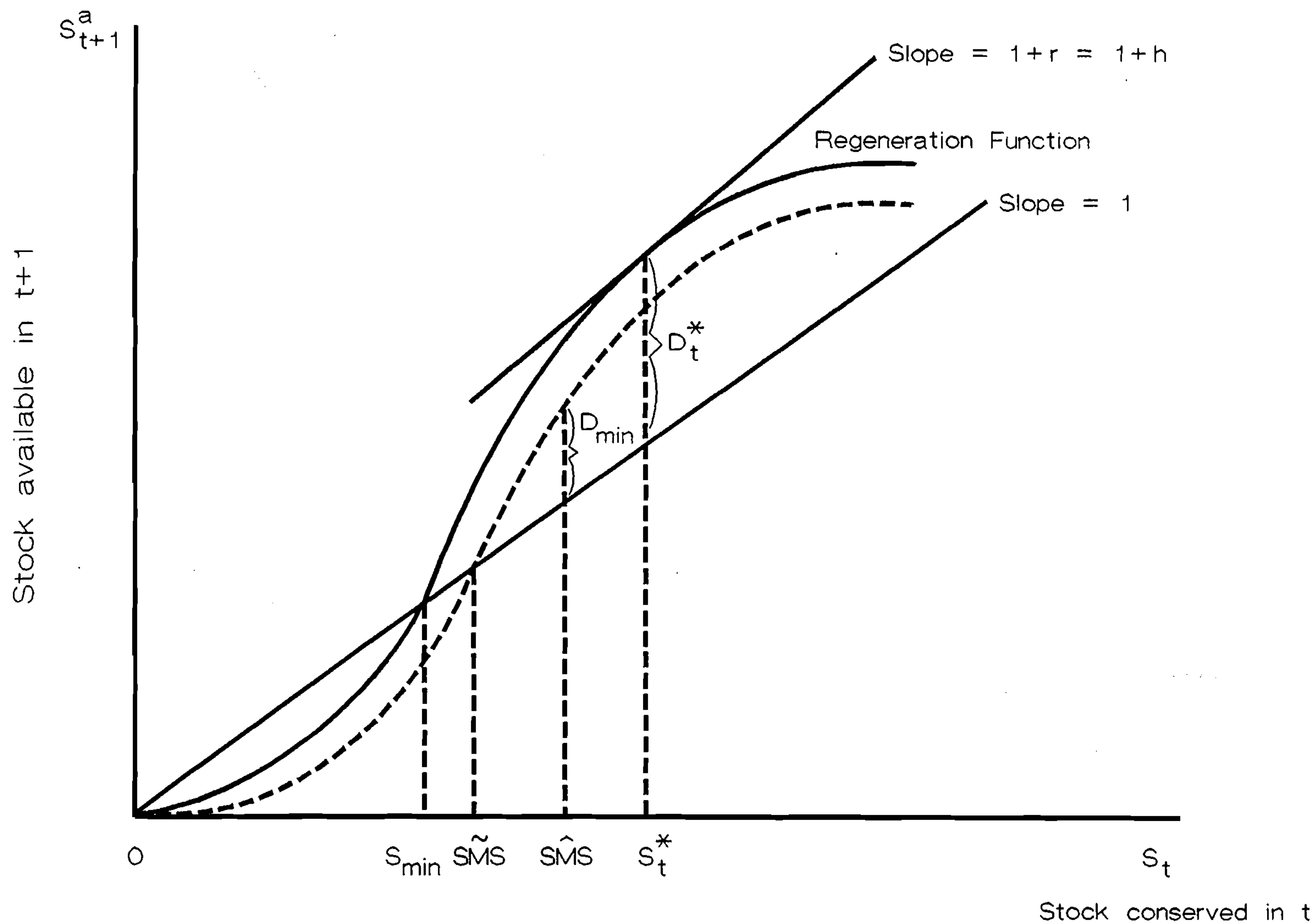


Figure. Setting the SMS.

CHOOSING ECONOMIC CONCEPTS FOR APPLICATION TO ENDANGERED SPECIES DECISIONS

Daniel D. Huppert, Associate Professor
School of Marine Affairs, HF-05
University of Washington
Seattle, Washington 98195

ABSTRACT Endangered species decisions occur in two stages: (1) the selection of species to preserve; and (2) selection of conservation measures. This paper summarizes the economics theory of species preservation, focusing on Bishop's Safe Minimum Standard (SMS) approach. The SMS approach deals only with the first stage decision, calling for preservation of all species unless the social costs are unacceptably large. Other approaches (e.g. Smith and Krutilla and Montgomery and Brown) place species conservation in an incremental decision context appropriate to the second stage. These approaches note that additional protection, incurring higher costs, yields a higher probabilities of species survival. The Endangered Species Act has an objective analogous to the SMS approach, but invokes decision procedures after the "listing decision" that admit incremental approaches and economic reasoning. Selection of critical habitat areas and of conservation measures in that habitat involve consideration of alternative actions and economic trade-offs. We conclude that: (1) assessment of economic cost for species protection is required by the ESA and is necessary for agency listing decisions under budget constraints; (2) evaluation of the cost-effectiveness of alternative conservation measures could improve efficiency of ESA decisions; and (3) benefit-cost analysis of species preservation decisions is not called for under the ESA, but benefits assessment on a case-by-case basis could guide use of the ESA's exemption provision.

Three stocks of Snake river salmon are listed as threatened or endangered species, and recovery planning is underway. Because these species live in an extensively developed river system, preservation actions could lead to sacrifice of some significant economic development benefits. Outputs of hydro power dams, agricultural irrigation, river transportation, Federal timber harvests, and Federal grazing activities could be curtailed in recovery efforts. Benefits of preserving particular stocks of salmon should be weighed against incremental economic development benefits to gauge reasonable total costs of preservation programs. Commercial and recreational fishery values, however, suggest that the fish runs are of lower economic value than development benefits. Existence values have not been estimated. Cost-effectiveness analysis of alternative recovery actions is another role for economics. Unfortunately, scientific evidence regarding effectiveness of alternatives (augmentation of Snake river flow through releases from up-stream storage reservoirs and drawdown of reservoirs at the four Federal dams) is highly contentious. Biological advice to decision makers has become polarized.

In response to this scientific uncertainty, the Salmon Recovery Team has outlined a strategy combining short-term remedial actions (e.g. captive broodstock programs and limitations on harvests) and experiments to ascertain which river operations patterns will be most beneficial. This delays key decisions until more reliable information is available, and it sets out specific actions to obtain better information. Labeled "adaptive management" by Kai Lee, this approach requires decision makers to admit to uncertainty, to exercise patience as knowledge accumulates on biological, not political time scales, and to devote agency resources to the experimental efforts. An economic justification for this solution seems straightforward, if not yet fully developed.

CHOOSING ECONOMIC CONCEPTS FOR APPLICATION TO ENDANGERED SPECIES DECISIONS

INTRODUCTION

Decreasing biological diversity, especially through species extinction, is an important economic problem whose resolution in public decision processes becomes increasingly contentious. The continuing debates over conservation of spotted owl and salmon habitats on Federal lands in the Pacific Northwest reveal the basic structure of the problem. Some well-established interest groups favor continued economic benefits associated with development and transformation of natural environments, while other interest groups focus on the loss of species diversity associated with development. In the spotted owl case, the forest products industry, associated communities, and consumers of wood products benefit from continued conversion of old-growth forests into even-aged stands of trees managed for sustained periodic cutting. The resulting "tree farms" display highly simplified ecological structures that eliminate some habitat types. Changes wrought in forest habitats will probably cause irreversible changes in ecosystem functions and species assemblages, including the extinction of species which are narrowly defined (i.e. sub-species, or reproductively isolated populations of broader taxonomic species) or which are highly adapted to the specific characteristics of old growth forests. While forest ecologists have designed "new forestry" techniques to save some characteristics of old-growth forests, modern high-yield forestry seems incompatible with preservation of the authentic ancient forest ecosystem. If this view is correct, an economic trade-off between expanded modern forestry and ecosystem preservation is unavoidable.

The public battle over management of old-growth forests may be overshadowed soon by analogous legal and political efforts to preserve Pacific salmon species in Pacific Northwest river basins which have been extensively modified by economic development. Numerous salmon stocks have already gone extinct, and many additional stocks are likely threatened with extinction (Nehlson, et al. 1991). In the Columbia River basin, total in-river salmon run sizes have dropped from an estimated level of 10 - 16 million fish in the late 1800s to a level of around

2 million. In recent years, around 80 percent of the Columbia river salmon catch has been of hatchery-produced rather than naturally spawning fish. Further, three species of salmon in the Columbia basin have been listed as threatened or endangered under the ESA by NMFS[1]:

(1) sockeye salmon (Oncorhynchus nerka) spawning in Redfish Lake within the Sawtooth National Recreation Area of Idaho,

(2) spring and summer chinook salmon (Oncorhynchus tshawytscha) spawning in tributaries and headwater habitats of the Snake River basin below Hell's canyon dam, and

(3) fall chinook salmon (Oncorhynchus tshawytscha) spawning in the mainstem of the Snake River between the Lower Granite dam reservoir and the Hell's canyon dam.

According to the Draft Snake River Salmon Recovery Plan Recommendations, total production of chinook salmon in the Snake river during the late 1800s was about 1.5 million fish. These salmon populations declined after construction of numerous dams on the Snake and Columbia Rivers.[2] During the 1980s returns of spring and summer chinook to the mid-Snake (above Lower Granite dam) averaged 9.6 thousand fish. Spawning runs of naturally spawning Snake river fall chinook fell to a low of 787 fish in 1990. The sockeye salmon returning to Redfish lake, never a

---

[1] The term "species" under the Endangered Species Act (ESA) differs from the common taxonomic definition. An ESA "species" is an isolated stock which interbreeds when mature, a subspecies, or distinct population segment. Salmon populations are naturally divided by spawning streams and timing of spawning run, with each local population possessing some genetic differences that are adaptive to local conditions. In practice, the number of salmon stocks (or "ESA species") identified depends upon how finely spawning seasons and river segments are divided. These divisions depend on the significance attached to particular genetic, physiological, or behavioral differences. Nehlson, et al. identified 214 "stocks" of natural Pacific salmon which are threatened to some degree, and 120 stocks that are already extinct. In his review of species definitions for salmon, Waples (1991) concludes that "distinct population segments" should be identified as ESA species if they are "Evolutionarily Significant Units" (ESUs), i.e., if they are effectively reproductively isolated and exhibit unique, evolutionarily important traits. The three listed Snake River species were determined to be ESUs.

[2] The Swan Falls dam on the upper Snaker river was the first to be constructed in 1901. Federal dam projects include four mainstem Columbia River dams (Bonneville, The Dalles, John Day, and McNary), and four lower Snake River dams (Lower Monumental, Ice Harbor, Little Goose, and Lower Granite). Idaho Power Company has three major dams on the mid-Snake river (Hell's Canyon, Oxbow, and Brownlee), and several smaller projects have blocked salmon runs on the Clearwater river, Salmon river, and Grand Ronde tributaries of the Snake river.

very large population, fell from an estimated 4 thousand fish in the 1950s and 1960 to as low as 1 fish in recent years. As shown in Table 1, the numerous causes for the decline of Columbia basin salmon include water diversions for agricultural irrigation, fishing in both river and ocean areas, and degradation of spawning and rearing habitats due to forestry, mining, and livestock grazing activities. While all of these economic activities contribute to the depletion of upstream stocks of salmon, the hydroelectric facilities are widely viewed as the major culprits. Given the strong prohibitions against taking of endangered species under the Endangered Species Act, the Pacific Northwest region may be forced soon to trade off the economic benefits of river development for preservation of biological diversity in the river ecosystem.

Commercial and recreational fishing values for salmon stocks are important, but the economic use values associated with salmon-depleting river development seem much larger. During 1986-90 the exvessel value of chinook salmon taken commercially in ocean fisheries off Washington, Oregon, British Columbia and southeast Alaska[3] averaged about $ 49 million. Commercial fishing in the Columbia river by treaty and non-treaty fishermen brought an average annual exvessel value of about $3 million. Using an estimated average net recreational value of $41.61 per fish (Olsen, Richards, and Scott, p. 53), the value for recreational chinook salmon harvests off Washington, Oregon, British Columbia, and southeast Alaska during 1986-90 was roughly $17 million. In contrast, hydroelectric power produced by the Federal dams in the Columbia river basin has a wholesale value in the neighborhood of $2.5 billion. While these numbers are not theoretically correct measures of economic benefits, and they are not cast in the appropriate marginal framework, they do provide some basis for gauging the relative importance of the fishery and non-fishery benefits. Even a ten-fold expansion of Columbia River chinook salmon runs would leave the economic value of salmon below that of the hydroelectric power system. Further, the river transportation, flood control, and agricultural irrigation projects dependent upon dams in the river are central to the economy of the Columbia river basin. It

[3] This is the region in which Snake river salmon are caught by commercial and recreational fisheries.

would seem reasonable, therefore, not to embark on a program to save all runs of salmon "at any cost".

It is more difficult to gauge the specific value of maintaining naturally-spawning salmon populations in the upper reaches of the Snake river. Apparently, preservation of naturally-spawning salmon stocks could contribute in three ways. First, the symbolic value of salmon to the Pacific Northwest culture and environment may be very great. This existence value for endangered salmon stocks could be estimated via the Contingent Value Method (CVM), but only preliminary and partial existence value information is yet available.[4] Second, the artificial propagation of salmon in hatcheries leads to in-breeding and genetic drift, making it necessary to infuse wild salmon genes into the hatchery system periodically. From this perspective the benefits of the hatchery program (e.g. the commercial and recreational use values) are dependent upon the preservation of the wild salmon runs. Third, long term survival of salmon species in the river system depend upon maintenance of the genetic diversity in wild stocks, as this diversity provides the material basis for adaptation to changing conditions.

This paper will focus on the use of economics as a source of information for the public decision processes concerning endangered species. We evaluate its contribution to two stages of decision making: (1) selection of species to preserve, and (2) selection of specific conservation and recovery measures for a selected species. Much of the published economics thinking about endangered species has focused on the first stage decision. As I describe below, the emphasis of this branch of thought is on the special conceptual problems posed by irreversibility of extinction, uncertainty regarding future use value, the magnitude of non-use values, and intergenerational equity. The second stage decision draws upon applied economics to guide effective implementation of existing legislation, especially the Endangered Species Act (ESA). The emphasis in this second stage is on economic assessment of alternative "critical habitat

[4] For example, Olsen, Richards, and Scott estimated the existence value for doubling all Columbia river basin salmon stocks.

designations" and recovery measures, using tools of benefit-cost analysis or cost-effectiveness analysis or simply economic cost analysis.

After briefly summarizing the economic concepts and theories from the literature, I offer some additional observations concerning effective use of economics in ESA decision processes. I will argue that, at least in the Snake River salmon case, uncertainty concerning the technical feasibility of proposed recovery measures is an important adjunct to uncertainty about future economic value of salmon sub-species. This additional uncertainty confounds both the first and second stage decision problems. When we cannot confidently assign costs of preservation, it is difficult to use cost-effectiveness analysis to develop a social cost function. Where the outcomes of proposed recovery measures are highly uncertain or seriously contested by the technical experts, we are left with only cost analysis. Second, I note that adaptive decisions under this type of uncertainty could include (1) short-term measures to preserve future options, linked with (2) deliberate experimentation and learning to dispel sources of ecological and economic uncertainty about recovery options. In this adaptive approach, the outcomes of earlier steps become prior information in later steps. Careful design of early decisions can promote learning to constructively guide better decisions in the future. The role of economics may be to help in the design of crucial experiments and to develop better information about non-market values.

ECONOMICS OF SPECIES PRESERVATION PROGRAMS

Space does not allow an extensive historical review of economic thought about endangered species. Instead, I provide a brief sketch of what seem to be the major elements of current, mainstream thought. This summary relies heavily on Bishop (1978, 1979, 1993), Smith and Krutilla (1979), Randall (1986, 1988), Tisdell (1990), and Pearce and Turner (1990). Most of this literature focuses on what I have called the first stage decision -- to preserve or not preserve a given species.

The Economics of the Preservation Choice

As originally developed by Bishop (1978) the policy choice involves a trade-off between economic development and species preservation. A special concern is the possibly large, but unknown future economic value of the endangered species. Bishop's papers reflect profound concern that we cannot adequately anticipate the functions that species perform in the ecosystem and that we cannot estimate the economic values we (or subsequent generations) might place on these species in the distant future. The inability to estimate future values is both structural (science may never accurately predict responses of complex systems and the future preferences are not knowable) and situational (we have not developed adequate empirical tools to estimate ecological functions and economic values)[5]. Added to this concern is the ethical problem that future generations are directly affected by the irreversible extinction, yet have no voice in the decision. In this situation, quantitative assessment via Benefit-Cost Analysis (BCA) of economic values lost through extinctions is a thin reed upon which to rest preservation policy. Bishop observes that "potential extinction creates an important public policy issue because there is little basis to judge which life-forms can be discarded without serious future social and economic consequences. To choose extinction creates the possibility of large future losses." (1978, p. 12) As a direct consequence of this, Bishop rejects standard BCA as a decision tool for endangered species. Instead, he proposes a modified version of Ciriacy-Wantrop's Safe Minimum Standard of conservation (SMS), calling for avoidance of irreversible losses (preservation of species) regardless of current estimates of economic costs and benefits.

Bishop and later writers on the SMS approach use a payoff matrix to illustrate the nature of the decision problem, and to show how it is equivalent to a strategy of minimizing the maximum potential loss (minimax loss). In Table 2, the second and third rows correspond to the dichotomous policy choice: (1) proceed with economic development which will cause extinction,

[5] Here Bishop makes the distinction between true uncertainty and risk. In the presence of risk, we proceed by assigning probabilities to alternative outcomes and then either choose the outcome with highest expected payoff or optimize a state-contingent utility function. With true uncertainty we cannot assign probabilities and outcomes with any confidence.

and (2) abandon or modify economic development to assure species preservation. Two alternate states of nature are captured by the second and third columns. State 1 represents the case in which the species being considered will ultimately be found to have no economic value to people. State 2 is the state in which the species is found to have significant economic value, represented as the "benefits of preservation" B_p. If we arbitrarily take the economic development option as the base case, continued benefits of development under State 1 involves an economic loss of zero because the benefits of preservation turn out to be zero. In State 2, the loss due to choosing extinction is B_p. The choice of the preservation option (row three) involves incurring some costs in the form of reduced benefits of economic development, represented as B_d. Under State 1, the preservation decision incurs a loss of B_d and there are no off-setting benefits of species preservation. Under State 2, the future value of the species is an offset to the costs of preservation, and the net social cost of preservation is B_d-B_p.

The choice of a safe minimum standard approach (that is, to choose preservation over economic development) is theoretically justified if (1) we cannot know whether State 1 or State 2 will prevail, (2) the potential future loss to society of extinguishing a species is likely to be large relative to the value of the economic development, and (3) lacking a basis for assigning probabilities to States 1 and 2, we follow a minimax loss strategy. Hence, if $B_d < B_p$, we choose option 2 in order to minimize the maximum possible loss. One objection to the SMS approach is that we do not know that the loss of economic development benefits is relatively small. Bishop addresses this objection by offering the modified SMS which calls for preservation of species, unless the social costs of preservation are <u>unacceptably large</u>. This proviso allows for the consideration of social costs, but Bishop notes that and economic calculation cannot determine what is unacceptably large. There must be some collective decision process to determine when costs are too large to accept. Another objection is that the adoption of this approach in practice may place too much emphasis on extreme outcomes (Brown, 1990); if most development actions do not necessarily result in extinction, it may be unreasonable to foreclose economic development options in order to avoid a highly unlikely event.

Randall (1986) provides another interpretation of the SMS approach. He suggests that the decision maker maximize the number of species saved subject to the budget constraint. "All species would be treated as having a positive but unknown expected value; implicitly all would be treated as equally valuable. Priorities would be set on the basis of opportunity costs; preserve that package which includes the most species given the cost constraint" (1986, p. 103). A further elaboration of this approach might call for assessment of each species' contribution to "biodiversity" and a selection of species to save which maximizes the amount of biodiversity preserved.

The various authors assume different objectives, decision-making procedures, facts concerning irreversibility and preservation costs and benefits, and level of uncertainty concerning outcomes of preservation and development efforts. Most of the attention is placed upon the uncertainty concerning losses attending the extinction of a species. An example of such potential loss is provided by the rosy periwinkle, which became valuable when it was found to yield alkaloids with medicinal uses. If we do not know everything about the species in question, nor what technological advances and scientific discoveries will reveal about new uses for the species, then we can imagine almost any species having extraordinary economic value in the future. Intergenerational equity questions can also be important in contemplating extinctions. Although the current generation may place minimal value on a species facing extinction, we do not know that future generations will do so; yet our decision to permit extinction binds the future generations to that outcome. Maintaining some minimal viable population of the species preserves future options. It seems that the prospect of future economic values becoming large either through scientific discovery or through changes in preferences, can motivate adoption of the Safe Minimum Standard of conservation approach.

In contrast to the SMS approach, the "Resources for the Future approach" developed by Krutilla and Fisher (1975) and Smith and Krutilla (1979) adopts "the perspective of conventional benefit-cost utilitarian framework ."[6] The RFF approach calls for a full enumeration of all benefits and costs of investment in a discrete project which involves a unique site whose modification will be irreversible. "The RFF decision rules can be derived from a statement of society's objective function, namely to maximize the discounted net benefits associated with the expenditures involved in a given project, together with an explicit treatment of the irreversibility inherent in the decision."[7] Smith and Krutilla (S-K) respecify the extinction problem as one which involves an uncertain critical zone beyond which development will cause destruction of a natural resource. S-K designate $B_d[S(t)]$ as the benefits from holding stock S(t) of the natural asset in a development state. Similarly, $B_p[S(t)]$ is the preservation benefit associated with a stock level S(t), and I(t) is the investment cost of the development project. Assume that the development costs and physical quantity of the stock are linearly related as

$$\frac{\mathbf{dS}}{\mathbf{dt}} = \sigma\mathbf{I(t)n} \text{ and } S(t) > 0$$

Assume that the critical zone below which the natural asset is irreversibly lost, $\bar{S}$, is a random variable distributed over the interval $(0,\infty)$ with probability density $g(\bar{S})$. Then any incremental increase in development generates direct development benefits, reduced benefits derived from the resources stock, and increased likelihood of extinction.

Smith and Krutilla describe a decision rule that maximizes the expected value of discounted net benefits given the probability of irreversible affects on the natural stock. The optimizing condition states that the investment in development proceed so long as the marginal benefits exceed the direct marginal costs plus the expected foregone marginal preservation benefits plus the marginal costs associated with an enhanced likelihood of extinction. S-K note that their approach differs from Bishop's SMS approach., as it is derived from an explicit

---

[6]Smith and Krutilla (1979) p. 372
[7]Ibid. p. 372-73.

objective function with direct account taken of the potential irreversibility. This approach treats those cases in which a stock of an endangered species may be reduced without necessarily leading to the extinction of that species. Hence, the S-K approach considers optimal investment decisions in the presence of uncertain thresholds on the irreversibility of development investments. Some modifications to the natural environment may be technically irreversible (cannot be restored to the exact, authentic natural state), but may not necessarily imply that the critical elements responsible for all preservation benefits are lost. Moreover, we may not know in advance whether these benefits will be lost. This important interaction between irreversibility and uncertainty was already exposed in Bishop's 1979 paper.

In Bishop's reply to S-K, he distinguishes between public decision making with risk, i.e. where alternative outcomes of investment decisions occur with known probability and known payoff, and decision-making under true uncertainty, i.e. where the probabilities of alternative outcomes are unknown. He claims that the SMS approach more accurately characterizes the uncertainty involved in species extinction. Because the S-K approach transforms the problem to one involving only risk, it assumes away a large part of the problem. Bishop's SMS approach treats uncertainty about impacts of extinction as pure uncertainty, which motivates the modified minimax principle. Further, Bishop notes that since the choice problem involves a value judgement about intergenerational equity, standard economic criteria cannot determine when the social costs are unacceptably large.

A view closer to the S-K approach is expressed by Montgomery and Brown (1992). They note that presentation of all-or-nothing choices (preservation or extinction) gives policy makers little scope for comparing alternative plans. More importantly, the all-or-nothing approach "fails to provide a basis for assessing the costs and gains of moving incrementally toward more or less species protection. The lack of a traditional marginal framework for reasoned analysis of trade-offs further aggravates the adversarial debate between advocates of species conservation and those likely to bear the cost." (p. 1) In their analysis of spotted owl preservation, Montgomery and Brown propose to develop a "supply curve of species survival" which displays the range of economic costs and associated survival probabilities available. They note that increasing the amounts of owl habitat conserved will improve the likelihood of owl survival. Nothing will achieve 100 percent assurance that spotted owls will survive forever. Hence, the analytical task is converted to one of linking specific volumes (and patterns) of preserved habitat to probability of survival, and to economic cost. With this information in hand, the decision makers can consider the trade-off of, say, another million dollars worth of sustained timber harvest versus an increase of some marginal percentage in owl survival probability.

In this view, the decision to list the species as endangered and to accept substantial curtailment of economic development benefits does not settle the issue as to whether preservation has been "chosen". Listing under the Endangered Species Act, for example, is simply a signal of determination to devote substantial effort to preservation. Many choices in species preservation will involve incremental commitments to improve survival chances. As more old-growth habitat is preserved and benefits of intensive forestry are sacrificed, the spotted owl's chance of survival is increased. This characteristic extends to the Snake River salmon case discussed below. In fact, the number of preservation options seem greater for salmon, and uncertainty about each option's ability to improve the chances of survival seems also to be greater. A major contribution of economics to endangered species decision making is to clearly identify the trade-offs between species preservation, economic costs, and wider measures of social cost.

J. C. Whitehead's (1992) paper on contingent valuation method for endangered species notes that contingent valuation mechanisms for endangered species have typically assumed certain supply, but have implicitly incorporated uncertain demand. That is, the survey respondent answers questions of willingness to pay (or accept compensation) by subjectively incorporating her own uncertainty concerning future demand. However, the respondent is apparently "paying for" certain preservation of the species. Whitehead argues that the survey design should incorporate uncertainty in both ex ante supply and demand. This insight mirrors the Montgomery and Brown formulation of the endangered species decision problem. We do not know how to preserve any species forever; and we have great uncertainty concerning some of the most celebrated preservation programs currently in operation (e.g. the California condor). The simple choice mechanism inherent in the typical CVM study abstracts from the real problems of preservation program implementation. An implication of the Montgomery and Brown view is that, even if the SMS approach is adopted in high-level species preservation decisions, program decisions will frequently hinge upon understanding the degree to which feasible policy options improve the chances of survival. Similarly, economic assessments of benefits of preservation should value the prospects for improved species survival rather than absolute survival versus extinction. [8]

The Endangered Species Act and Safe Minimum Standard of Conservation

The SMS approach closely resembles the approach developed by Congress in the Endangered Species Act of 1973 (plus 1978 amendments). The ESA process involves six stages: (1) the listing decision, (2) the designation of critical habitat, (3) jeopardy determinations (in which the Secretary of Interior or Commerce issues a "biological opinion" that a Federal agency program does or does not jeopardize an endangered species), (4) Section 7 interagency consultations (in which the action agency consults with the listing agency to avoid jeopardizing a species), (5) Section 7 exemption process, and (6) recovery planning and management. The extent to which economic factors can be considered is determined by the text of the Act, the legislative history of the Act, administrative discretion exercised by Federal agencies, and legal

[8] The author finds that a more recent paper by Ready and Bishop (1991) succinctly incorporates this notion of uncertainty in supply. They describe a pay-off matrix in which it uncertain that the preservation action will lead to any benefit, which they call the "lottery game". That paper also amends the earlier conclusion that SMS is consistent with a minimax loss strategy. Ready and Bishop show that the minimax loss rule can lead to development rather than preservation in the lottery game. The author regrets that Ready and Bishop's presentation was not incorporated in the main text of this paper.

actions initiated by public interest groups or environmental activists. As summarized in Table 3, economic information can be considered in (a) weighing the benefits of including an area in critical habitat against the benefits of excluding that area, (b) evaluating alternative agency actions to avoid adversely impacting a listed species or its habitat, (c) appealing for Section 7 exemption to the Endangered Species Committee, and (d) estimating the cost of recovery measures considered in the Recovery Plan.[9] As noted by Bishop (1993) and Ready and Bishop (1991), the ESA follows the SMS strategy of formally protecting all species. Economics has not and likely will not be important in naming critical habitats, because no specific action (and, hence, no specific economic consequences) are entailed in the critical habitat designation. On the other hand, the Section 7 exemption process is tantamount to the determination that social costs of species preservation are "unacceptably large". By a majority vote of at least five to seven, the Committee may grant an exemption, if they determine that:

> *(i) there are no reasonable and prudent alternatives to the...action; (ii) the benefits of such action clearly outweigh the benefits of alternative actions consistent with conserving the species or its critical habitat, and such action is in the public interest; (iii) the action is of regional and national significance; and (iv) neither the federal agency concerned nor the exemption applicant made any irreversible or irretrievable commitment of resources prohibited by subsection (d) of this section..*[10]

While economic costs are clearly a major factor in appeals for exemption from the God Committee, that process is rarely invoked. Since the 1978 ESA amendments created the exemption process, the Committee has voted on only three applications: the Tellico Dam, Graylocks Dam, and BLM timber sales in the Pacific Northwest. Exemptions for the Tellico and Graylocks Dams were denied. The BLM applied for exemption from ESA obligations for the sale of 44 tracts of timber. The Endangered Species Committee exempted 13 of the 44, denying

---

[9] See Gleaves and Wellman (1992) for a more extensive discussion of economics in the ESA process.
[10] ESA section 7 (h)(A).

exemption for 31 tracts. Several months later, however, the BLM withdrew its application for exemption, without having proceeded with the sale of the approved 13 tracts.

I suspect that the most important contribution of economics will be in assisting agency choice of actions to avoid jeopardy opinions and in species recovery plans. The broad mandate to develop recovery plans should include a wide-ranging evaluation of alternative strategies with net economic benefits being one criteria along with population recovery rates and species survival probability.

ECONOMICS OF PRESERVING SNAKE RIVER SALMON

The causes for decline in Snake River salmon abundance lead to proposed system changes to encourage species recovery. Since the system is complex and the causes of decline numerous, many different actions have been proposed, some of which are mutually exclusive. Selection of preservation and recovery methods should proceed on the basis of a comprehensive model of salmon life history. Roughly, survival from juvenile to spawning adult takes the fish through four sequential stages, each identified with specific habitats:

(1) Freshwater spawning and rearing habitat: spawning gravels of cold, flowing streams; growth through fry and fingerling stages to smoltification (physiological preparation for downstream migration to saltwater);

(2) Downstream Migration Corridor: migration of smolts downstream through free-flowing rivers, impoundments behind dams, and hydroelectric dams to the Columbia River estuary and Pacific ocean;

(3) North Pacific Ocean: feeding and growth during ocean migration phase;

(4) Upstream Migration Corridor: migration of mature adult salmon up through lower Columbia River, eight federal hydroelectric facilities equipped with fish ladders and free-flowing river to spawning sites.

Strategies to improve survival of these species have focused on those stages in which (a) especially high mortality rates are thought to occur, and (b) the system can be manipulated through federal agency programs. Table 4 lists some of the special problems faced by the endangered salmon species and some recovery actions being considered for each life stage. Generally, the protection and improvement of freshwater spawning and rearing habitats involves land management practices, especially on Federal lands in eastern Oregon and Idaho. Downstream migration through the eight Federal dams causes substantial mortality, but the magnitudes are hotly disputed. Present efforts to reduce this mortality include diversion devices (e.g. traveling screens) to direct the out-migrating smolts away from turbines, spilling of water over the spillway to carry smolts through safely, collection of smolts at Lower Granite Dam and other dams for transportation by barge to the Columbia river estuary, and harvesting of a predatory fish -- the northern squawfish. While 100s of million of dollars have been spent on these efforts, the downstream mortality is apparently still too high to stabilize the populations. During ocean migration and feeding, the salmon are subject to fishing mortality off Southeast Alaska, British Columbia, and the coasts of Oregon and Washington, and they are caught incidentally in the commercial, recreational and tribal fisheries in the Columbia river. Due to careful regulation of the fishery, fishing mortality seems unimportant for the spring and summer chinook, but fishing mortality can be as high as 60% for fall chinook. During upstream migration, passage through dams is facilitated by fish ladders at all dams up to Hell's canyon dam. While some improvements in these facilities are being proposed, they are apparently not the major source of mortality.

While accurate estimates of cost cannot yet be attached to most proposed Snake river salmon recovery options, order-of-magnitude estimates are possible. For example, curtailment of fisheries to reduce mortality of adult fish seems to be relatively inexpensive, costing at most tens of millions of dollars per year. Salmon-related restrictions on timber harvests, grazing activities, and recreation on national forests in the Snake river drainage also seem relatively inexpensive.

Estimates by the U. S. Forest Service summarized in Huppert, Fluharty and Kenney attach costs of $ 2.4 million/year to livestock grazing measures, $ 220 thousand/year to timber-related costs, and $ 1.6 - 2.1 million/year to reduced recreation. More expensive are the various improvements to the downstream fish facilities, including juvenile fish bypass facilities at Federal dam and barge transportation systems for smolts. The measure likely to entail the greatest cost is the adoption of full drawdown of Snake river reservoirs to below current spillway height. To reduce reservoirs behind the dams to near river level would require substantial engineering and construction work taking 10 - 15 years and probably costing more than a billion dollars overall.

Estimated costs of associated with changing the river operation hinge upon assumptions regarding demand elasticity for final products (especially price elasticity of electricity demand), availability of substitute services (e.g. rail transport to replace river barge shipments), and economic institutions and procedures of the salmon recovery implementation effort. That these factors can be extremely important is not surprising to economists. For example, if water markets existed, river flow augmentation could proceed by purchase of water rights from those having lowest Willingness to Accept Compensation (WTA) for reductions in out-of-stream water use (see Huppert, Fluharty, and Kenney; p. 3-52). Functioning water markets could distribute the reduced water use in an efficient fashion. Other procedures to reduce water use, such as proportional reduction of all water rights or curtailment of water rights dependent upon Federal storage projects, could impose much larger producer surplus losses. Similarly, reduced fishing could be accomplished at lower overall economic cost if accomplished through purchase of tradeable individual quotas rather than arbitrary curtailment of segments of the fishery.

If we had reliable estimates of economic cost associated with the recovery options, we would have a pay-off matrix which is much more complex that the two-by-two matrix used to illustrate the SMS approach. It would look something like the matrix in Table 5, where the first column lists recovery options, and the second column lists economic development losses associated with each option (assuming that the development and extinction option is the status quo). The terms $-\Delta_i B_d$ represent the change in economic development benefits for option i. Further, with reliable estimates of recovery effectiveness, we could re-rank the options, placing the most effective per dollar of cost at the top and dropping the less effective towards the bottom of the list. To obtain the most recovery for a given total cost we would select from the top. The greater the budget, the further down the list we would go. While the decision to devote a large budget to the preservation effort may be largely social rather than economic, the allocation of that budget could logically follow a cost-effectiveness procedure. The measure of effectiveness could be the likelihood of species survival, the size of the population increase, or a partial measure such as improved survival in downstream migration phase. This sort of cost effectiveness analysis was proposed by Olsen (1993) for the Snake river salmon, and by Paulsen, Hyman, and Wernstedt (1993) for numerous Columbia river basin salmon populations.

In the Snake River salmon case, however, uncertainty pervades the analysis of biological effectiveness of alternative preservation actions even more than the economic cost estimates. Although a variety of conservation actions have been proposed and examined over the past decade and a half, the Salmon Recovery Team finds little reliable scientific evidence that some of the major options would be effective. The uncertainty is not simply about precision of estimated effects. Rather, there are hotly contested issues concerning drawdown of reservoirs to spillway height. Some experts, including the Salmon Recovery Team, argue that this would not be even measurably effective. Others argue that it is the single most important step for recovery. Transportation of smolts is another polarizing issue. Competent biologists take diametrically opposed positions concerning the effectiveness of such actions. Hence, for now at least the usual cost effectiveness approach is defeated by the extent and character of the uncertainties.

Other approaches to decision making could incorporate the notions of "precautionary actions" and adaptive management. The precautionary agenda would include actions which "buy time", even though they do not promise long-term survival of the species in question. Captive broodstock programs, as currently implemented by the National Marine Fisheries Service for

Snake river sockeye salmon, maintain viable populations of endangered salmon until problems with the spawning habitat and downstream migration survival are solved. The "adapative management" approach redirects attention from the strict "all-or-nothing" character of species preservation decisions, and toward the process of learning and resolving uncertainties in an incremental fashion. As stated succinctly by Kai Lee "adaptive management is an approach to natural resource policy that embodies a simple imperative: policies are experiments; *learn from them.*" (p. 9) With adaptive management, we would engage in a sequence of policy actions, starting with those designed to preserve species while we attempt to learn more (a) about the nature of the genetic diversity embodied in the species; (b) the potential future economic value of the species; and (c) the feasibility, effectiveness, and costs of a range of possible preservation approaches. For a simplistic example, husbanding a captive broodstock of the endangered salmon during the first decade after listing of the species, preserves opportunities to re-introduce the fish into habitats after they have been altered to improve survival probabilities. This approach draws us away from the either/or nature of the current debate over ESA decisions, and invites us to develop innovative combinations of approaches which can be "tested" over time in a deliberate fashion. Presumably, questions amenable to biological/ecological research and to economic research can be researched during the initial phase of a preservation program in which irreversibilties are avoided through incremental decision making. Delaying the ultimate preservation/extinction decision in this fashion, under some circumstances, will provide time for human knowledge to improve, so that better decisions can be made later. In a practical sense, a focused program of experimentation and research might provide substantially more information about the most important variables to which substantial great uncertainty adheres. This adaptive approach could give way to a more conventional benefit-cost analysis or to another SMS approach, depending upon whether the amount of learning is adequate to resolve the main questions.

I find that an incremental approach of this sort is helpful whether the decision at hand is a preservation decision, or (as under the ESA) only the narrower choice of preservation method is considered. Before concluding that adaptive management is a solution to the problems of uncertainty, I must offer a warning. The shift to incremental decisions, with each subsequent decision based upon accumulated knowledge, could be something of a sleight-of-hand, distracting attention away from the fundamental issue of preservation and into the details of research and learning. No matter how effective the adaptive approach, we must ultimately face the question of whether the expected social costs of continued preservation are balanced by prospective future values. Another important research agenda, therefore, is the improvement of our ability to assess the value of the increased biodiversity inherent in saving additional strains of salmon species. A first step in that assessment must be the measurement of biodiversity itself. Solow, Polasky, and Broadus (1993) have examined a constructive approach to measuring genetic biodiversity, using "genetic distance" as a metric. Upon further development, this technique could help us to assess, for example, whether saving the Snake river chinook adds a significant amount of biodiversity to that already incorporated in other Columbia river chinook populations. This would be an important advancement. Economic assessment of endangered species strategy in the Columbia river will ultimately rely upon (a) improved measures of biodiversity and value of diversity, (b) improved ability to predict the effects of conservation actions, and (c) improved measurement of economic development losses associated with the conservation alternatives.

References

Bishop, Richard C. 1978. "Endangered species and uncertainty: the economics of a safe minimum standard," American Journal of Agricultural Economics, 60(1):10-18.

Bishop, Richard C. 1979. Endangered species, irreversibilities, and uncertainty: A Reply. American Journal of Agricultural Economics 61(2):376-379.

Bishop, Richard C. 1993. Economic efficiency, sustainability, and biodiversity. AMBIO, 22(2-3):69-73.

Brown, Garner M., Jr. 1990. Valuation of Genetic Resources. Chap. 5 in Orians, Gordon H., Gardner M. Brown, Jr., William E. Kunin, Joseph E. Swierzbinski. The Reservation and Valuation of Biological Resources. Seattle: University of Washington Press.

General Accounting Office (GAO). 1993. Endangered Species, Potential Economic Costs of Further Protection for Columbia River Salmon. GAO/RCED-93-41. Washington, D.C.

Gleaves, Karl and Katharine Wellman. 1992. Economics and the Endangered Species act. *The Public Land Law Review* 13: 149-165.

Huppert, Daniel D., David L. Fluharty, and Elizabeth S. Kenney, Economic Effects of Management Measures Within the Range of Potential Critical Habitat for Snake River Endangered and Threatened Salmon Species, Seattle, University of Washington School of Marine Affairs, June 4, 1992.

Kenney, Elizabeth. 1993. Economics and the Endangered Species Act: The Snake River Salmon Cast Study. Unpublished Master's Thesis, School of Marine Affairs, University of Washington. Seattle, WA.

Krutilla, John V. and Anthony C. Fisher. 1975. The Economics of Natural Environments: Studies in the Valuation of Commodity and Amenity Resources. Baltimore: Johns Hopkins University Press.

Lee, Kai N. 1993. Compass and Gyroscope, Integrating Science and Politics for the Environment. Washington D. C.: Island Press.

Montgomery, Claire and Gardner M. Brown. 1992. Economics of species preservation: the spotted owl case. Contemporary Policy Issues X(April): 1- 12.

Nehlsen, Willa, Jack E. Williams, and James A. Lichatowich. 1991. Pacific Salmon at the Crossroads: Stocks at Risk from California, Oregon, Idaho, and Washington. Fisheries, A Bulletin of the American Fisheries, Vol. 16, No. 2. pp 4-21.

Olsen, D., J. Richards, and R. D. Scott. 1991. "Existence and Sport Values for Doubling the Size of Columbia River Basin Salmon and Steelhead Runs". RIVERS 2(1):44-56.

Olsen, Darryll. 1992. The Path Toward Cost-Effectiveness. A working paper analyzing the cost-effectiveness of major salmon recovery alternatives. Northwest Irrigation Utilities, Portland, Oregon.

Paulsen, Charles M., Jeffrey B. Hyman, and Kris Wernstedt. 1993. Above Bonneville Passage and Propagation Cost Effectiveness Analysis. U.S. Department of Energy, Bonneville Power Administration Project Number 85-087-01. Portland, Oregon.

Randall, Alan. 1986. Human preferences, economics and the preservation of species. In B. G. Gordon (Editor), The Preservation of Species: The Value of Biological Diversity. Princeton University Press, Princeton, NJ. pp 79-109.

Randall, Alan. 1988. What mainstream economists have to say about the value of biodiversity. In E. O. Wilson, Editor Biodiversity, Washington D.C.: National Academy Press, Chap. 25 pp 217-223.

Ready, Richard and Richard Bishop. 1991. Endangered species and the safe minimum standard. American Journal of Agricultural Economics 73(2):309-312.

Smith, V. Kerry and John V. Krutilla. 1979. Endangered species, irreversibilities, and uncertainty: A Comment. American Journal of Agricultural Economics 61(2):371-375.

Snake River Salmon Recovery Team. October 1993. Draft Snake River Salmon Recovery Plan Recommendations. c/o National Marine Fisheries Service, Portland, OR.

Solow, Andrew, Stephen Polasky, and James Broadus. 1993. On the measurement of biological diversity. Journal of Environmental Economics and Management 24(1):60-68.

Tisdell, Clem, "Economics and the Debate About Preservation of Species, Crop Varieties and Genetic Diversity", Ecological Economics, Amsterdam, Elsevier Science Publishers B. V., 1990, 2, pp. 77-90.

Walters, Carl, Adaptive Management of Renewable Resources, New York, McMillan Publishing Co., 1986.

Waples, Robin S., Definition of "Species" Under the Endangered Species Act: Application to Pacific Salmon, NOAA Technical Memorandum NMFS F/NWC-194, March 1991, p. v.

Whitehead, John C. 1992. *Ex ante* willingness to pay with supply and demand uncertainty: implications for valuing a sea turtle protection program. Applied Economics 24:981-988.

Whittlesey, Norman K. and Philip R. Wandschneider. 1992. Salmon Recovery: As Viewed by Two Economists. *Choices* 7(2): 3 - 5.

Yaffee, Steven L., Prohibitive Policy - Implementing the Federal Endangered Species Act, Cambridge, MA, MIT Press, 1982, pp. 1-162.

Table 1. **NMFS' "Causes for Decline" of Snake River Salmon**

- **Destruction of Habitat or Curtailment of Species' Range**
  - Hydropower system: Slows freshwater migration, floods spawning grounds,
  - Timber harvest: Can degrade spawning areas by increased siltation, higher water temp.
  - Mining: Acid mine wastes degrade water quality, tailings may damage spawning rounds
  - Grazing: cattle often trample stream banks, pollute streams, disturb spawning gravels.

- **Overutilization**
  - Commercial and Recreational Harvests in River
  - U. S. Commercial Ocean Fisheries

- **Disease or Predation**
  - Freshwater Predation (squawfish in reservoirs, seals and sea lions in lower Col. R.)
  - Freshwater Disease

- **Inadequacy of Existing Regulatory Mechanisms**
  - Fish and Wildlife Coordination Act
  - Federal Power Act and Mitchell Act
  - Pacific Northwest Electric Power Planning & Conservation Act

- **Other Natural Factors**
  - Droughts: water flows below normal during recent years
  - El Nino: reduces ocean success, growth, & survival during ocean migration phase

---

Sources. National Marine Fisheries Service, Environmental and Technical Service Division. 1991. Factors for Decline, A Supplement to the Notice of Determination for Snake River Spring/Summer Chinook Salmon under the Endangered Species Act. Portland, Oregon.

__________. 1991. Factors for Decline, A Supplement to the Notice of Determination for Snake River Fall Chinook Salmon under the Endangered Species Act. Portland, Oregon.

Table 2. Pay-off Matrix for an Endangered Species Decision

	State 1 -Species Not Valuable	**State 2** - Species Valuable	**Maximum Loss**
Extinction	0	B_p	B_p
Preservation	B_d	B_d-B_p	B_d

Based upon Bishop (1978)

Table 3. Summary of ESA Steps and Economic Contribution to Decisions

STEPS in ESA Decision Process	Scope for Economics	Apparent Importance of Economics in Decisions	Economic Concepts or Analytical Method
1. Listing Decision	None officially. *Budgetary limits slow consideration of listings*	None, *but agency may seek to list as many species as possible with given budget*	*Maximize subject to budget constraint; requires direct cost analysis.*
2. Critical Habitat Designation	Consideration of economic impact. Weigh benefits of including an area against benefits of excluding an area	Broad prohibitions on "taking" make this less importance than ESA language suggests.	Techniques for quantifying costs and benefits applied to additional restrictions on use of habitat
3. Section 7 - Findings of Jeopardy or No-Jeopardy	None - exclusively a biological/ecological assessment	None	None
4. Section 7 - Formulating Alternatives to Avoid Jeopardy	Agencies seek to comply with ESA while minimizing loss in services delivered to constituents	This is a very active area of activity under Federal ESA administration.	Main method is cost analysis and cost-effectiveness
5. Exemption from No-Jeopardy Mandate (God Committee)	Explicit consideration of substantial economic loses due to Agency compliance.	Economic assessment would seem to be an integral element of case for exemption.	Economic cost and "impact" analysis are particularly relevant.
6. Recovery Planning	Explicit call for "time and cost" assessment; weighing of economic consequences in planning .	Economic evaluation of alternative approaches could be extremely useful, subject to biological uncertainties.	Full suite of cost and benefit evaluation tools organized in a Cost Effectiveness analysis.

Table 4. Summary of Problems at Different Life Stages, and Some Suggested Solutions.

LIFE STAGE	SPECIAL PROBLEMS	SOME SOLUTIONS
Freshwater Spawning and Rearing	Spawning grounds inaccessible, flooded by reservoirs, or degraded by natural resource industries	Control timber harvest, mining, and grazing in anadromous fish habitat
	Excess competition for food due to hatchery practices	Reduce hatchery take of wild fish for brood stock, reduce releases of fish in
Downstream Migration	Progress slowed by reservoirs, predation at dams, mortality in turbines and bypass systems	1. Augment river flows through releases of stored water 2. Draw down reservoirs to reduce cross-sectional area and increase current speed. 3. Collect and transport smolts downstream by barge 4. Improve passage facilities to reduce mortality at dams, spill water over spillways. 5. Harvest predators
Ocean	1. Ocean occasionally poor for salmon growth and survival: (El Nino)	1. No options here.
	2. Commercial fishing off U. S. and Canada takes significant numbers of Fall Chinook.	2. Reduce harvest rates in ocean fisheries affecting Snake River salmon. This includes Canadian fishery off Vancouver Island.
Upstream Migration	1. Sport and Commercial harvest by Treaty and non-treaty fishermen.	1. Buy-out river gill net licenses; invent new gear to capture fish live, allowing sorting; move treaty fishery upstream of Snake R confluence
	2. Passage through dams at fish ladders; low flows cause increased water temperature in lower Snake R.	2. Improve fish ladders; release additional water during late summer migration period to reduce water temperature in lower Snake.

Table 5. Pay-off Matrix with Uncertain Recovery Actions

	State 1	State 2	Probability of Extinction
Status Quo (Extinction)	0	$-B_p$	100 %
Action 1	$-D_1B_d$	?	? %
Action 2	$-D_2B_d$	?	? %
•		?	? %
•			
•			
Action N	$-D_NB_d$	?	? %

USE OF ECONOMICS TO IDENTIFY COSTS AND INSTITUTIONAL MECHANISMS FOR ADJUSTMENT TO SALMON RECOVERY IN THE PACIFIC NORTHWEST

Joel R. Hamilton
Dept of Agr. Economics
University of Idaho

INTRODUCTION

Soon after the first petitions were filed asking that several stocks of Columbia and Snake River salmon be listed as endangered, several of us received invitations to attend a meeting called by the Agricultural Experiment Station Directors of Oregon, Idaho, and Washington to explore what actions the three Land Grant Universities, WSU. OSU, and UI, should take in response to the impending listing. At that meeting we formed what we called the Universities Task Force on Salmon and the Columbia River System. This loosely organized group consists mostly of economists from the three Land Grant Universities, plus several fisheries biologists (including one from the University of Washington). We set two main goals for the task force, to help educate the public about salmon issues, and to help make the expertise of the universities available to those making decisions about whether to list salmon as endangered and what recovery actions to take.

The task force was then faced with meeting that educational and decision support role with no financial support except for time and travel money bootlegged from our other research activities. We were certainly not positioned to assemble a salmon "think tank" in competition with the other players. The involved agencies (Bonneville Power Administration, US Army Corps of Engineers, Bureau of Reclamation, the National Marine Fisheries Service, the Northwest Power Planning Council, etc.) all had existing research arms which soon cranked up to address the salmon issue. Similarly the various interest groups (irrigators, the Direct Service Industries, users of barge transportation, recreation and commercial fishing interests, etc.) soon assembled their research groups, consultants, and public relations experts.

In this environment, the members of the task force have tried generally to be big picture people -- we have attempted to keep the whole process in view, the legal environment, the analytic process, the problems, the likely impacts on various constituencies, and the lessons which we can learn from salmon which might apply to other endangered species. We have also tried to be guardians of good science. Most of us on the task force have found ourselves reviewing research reports and public statements from various participants, and trying to pass academic judgement on what we see. Many of the economist members of the task force participated as members of the NMFS Economics Technical Committee, which Dan Huppert has already told you about.

What I want to talk about today are several interesting economics-related conceptual and theoretical issues which I have encountered as I have tried to play this role.

WITH-WITHOUT PARADIGM

One of the great difficulties of trying to do economic analysis in an area as complex as salmon recovery is the difficulty of deciding exactly which proposed recovery actions should be the focus of the analysis. It is helpful to view the measurement of the economic effects of salmon recovery as a two step process. The first step is to properly identify the future equilibrium conditions that would be likely to exist "without" a salmon recovery program. The next step is to estimate the alternative future equilibrium that is likely "with" a salmon recovery plan. It is the comparison of these two dynamic equilibria which provides the framework for evaluating the types, magnitudes, and duration of changes that can be judged as economic costs and benefits from the recovery process.

Unfortunately this with-without paradigm is difficult to apply cleanly to salmon recovery. Efforts toward salmon recovery are not solely the product of the endangered species act. Various regional resource agencies have maintained hatchery, fish transportation, dam modification, and harvest management programs for many years. One of the main reasons for the passage of the Northwest Power Planning act a decade ago was to give anadromous fish a higher profile in regional power planning. Even without the recent push from the Endangered Species Act, a large and growing amount of effort and money would have been devoted to maintaining and recovering regional salmon runs.

It is difficult to define with and without scenarios which frame the analytic question clearly. Should we be interested in estimating the economic effects of only that increment of recovery actions prompted by the ESA listing of several salmon stocks as endangered? Should we be looking at all recovery actions, even those which would probably have gone on without the ESA listing? What actions would various agencies have implemented if the ESA listing hadn't happened? Did the ESA listing serve as a "club" to force other agencies to pay more attention to their obligations to salmon?

Probably the most difficult part of trying to conceptualize with-without scenarios for salmon recovery actions is to properly describe the fate of the salmon under each scenario. I have spent many years feeling that maybe my hard science friends were on to something when they denigrate economics as fuzzy and lacking predictive power. Now I have found a science discipline that is even fuzzier than economics -- fisheries biology. These biologists have so far shown little unanimity in projecting the effects of proposed recovery actions, or the effects of not taking these actions. This shouldn't be too surprising, since the actions now being proposed are often outside the range of the data of past experimental work, or even outside of past conceptual frameworks used in fishery science.

IMPACTS VERSUS BENEFITS & COSTS

Even if one succeeds in defining the physical, biological and institutional scenarios to be used in economic analysis, there remain many other conceptual problems. One of the most troublesome problems is the difficulty of moving from estimated impacts to the harder to estimate measures of cost or benefit. Efforts to achieve the recovery of threatened and endangered salmon stocks in the Columbia River basin will have many economic effects on the regional economy. There will be both gainers and losers resulting from the changes in resource management associated with the recovery efforts. Many economic interests are built around uses of the river resources that comprise salmon habitat.

Direct impacts are the initial changes in output and income from these directly affected sectors brought about by some recovery action. Examples are the changes in farm income caused by restricted access to barge transport of grain, the change in spending of sport fishermen caused by restrictions on harvest, the changes in agency spending needed to plan and implement the drawdown experiment, or the changes in value of hydropower production caused by drawdown implementation.

Secondary impacts flow from the primary ones, as the initial impacts work their way through the rest of the economy. If farmers are affected by drawdown, they will reduce their spending for both production inputs and personal consumption. If sport fishermen reduce their spending, this hurts those who make their living supplying fishing gear and guide services. If agencies change their spending, this too works its way through the local and regional economy.

As the changes in resource use required for salmon recovery are carried out, affected economic sectors will adjust to a new equilibrium. If these sectors have adequate knowledge of what the new system will look like, and adequate time to make appropriate investments and managerial changes, most economic sectors will adjust to the new equilibrium with welfare returns not unlike those enjoyed at the old equilibrium. Some differences (good and bad) may persist and be considered as a permanent change in welfare.

Those who participate in discussions about salmon recovery (including economists and politicians) tend to focus disproportionately on economic sectors which may suffer damage. These are the constituencies who demand that their problems be heard. It is important to recognize that recovery actions will help some while hurting others. While river drawdown may hurt barge operators and farmers, it might also help truckers and railroads. Recovery actions (and especially successful recovery) will help some recreation sectors while hurting others. Even the loss of hydropower generation associated with recovery actions will probably benefit some sectors. The challenge is to think globally enough so that both the "with" and "without" scenarios encompass the full range of positive and negative effects.

There are two main conceptual problems, "displacement" and "reemployment", which complicate the interpretation of the primary and secondary impacts of salmon recovery actions. Displacement results when the primary impact on one economic sector is accompanied by an offsetting impact on some other sector. For example, if drawdown damages flatwater recreation, people will not cease to recreate and certainly will not bank the money they would have spent in the losing sector. They will quickly turn to other forms of recreation, or other consumption spending, probably in the same general region. Similarly, if agencies increase their spending on salmon recovery activities, this may be just a reallocation from other programs.

Reemployment results when factors of production are initially idled by a recovery action, but through time find alternative employment. For example, a commercial fishery worker might lose his job because of harvest restrictions. After a period of unemployment he might be reemployed as a construction worker. Similarly, if farming income were to be severely damaged by drawdown, then those workers who depend on the spending of farm income would be forced to seek alternative employment. It helps to think in terms of the classic factors of production -- land, labor, and capital. Labor tends to be quite mobile. If one becomes unemployed, there is real incentive to diligently seek alternative employment. Capital tends to be less mobile. Money, operating capital, can be quite easily shifted to another enterprise. Physical capital, machinery, is also somewhat mobile, because it can be relocated or sold to another user. However fixed capital such as buildings, or other improvements to land are quite immobile. The salvage value for used building or a used canal may be minimal. Land is often the least mobile of the three. Obviously it can't be moved, and sometimes it has few alternative uses. Land presently used for a gas station may serve quite nicely as a site for a used car lot. On the other hand, land presently used for irrigated farming probably has little alternative use but low valued dryland grazing.

A part of this process, however, is the cost of adjusting to the new equilibrium. While some labor and capital will move to new occupations, some temporary unemployment or underemployment may occur. Some undepreciated capital may be abandoned as it becomes unusable, or as new technologies are adopted. While the new equilibrium position may provide welfare returns quite similar to the old position, transition to the new conditions will involve some very real costs. The task facing economists as they try to estimate the economic effects of salmon recovery, is to estimate the amount of permanent welfare change that may persist as well as the costs of reaching the new equilibrium position.

Measuring the permanent shifts in economic welfare is difficult. Because most mobile labor and capital will flow to new uses at values not unlike previous employment, idling them does not constitute a permanent loss. The time and effort required to reach the new employment positions will determine the cost of adjustment, but this is a one-time, rather than a permanent cost. Only if it can be demonstrated that unemployment will be long-term, or that future employment will be inferior to present, can long-term costs be claimed. Even here, the effects are not permanent, since ultimately all labor dies, and all physical capital is fully depreciated. On the other hand, salmon recovery will most likely involve a permanent loss in hydropower production, a cost that will continue for a very long time.

The bottom line is that many of the economic impacts of salmon recovery are short to intermediate term adjustment impacts -- not permanent impacts. The direct impacts will be very real, perhaps even large, and certainly very important to the people, industries, and communities involved, but most of the resources rendered idle will eventually find reemployment elsewhere, often at returns not especially inferior to their previous incomes. What really counts, is neither the primary nor the secondary costs per.se., but the adjustment costs (the incomes lost during adjustment, the cost of finding alternative employment, and perhaps an ongoing cost if the new employment is inferior). This logic is the reason why secondary effects are routinely excluded from national level benefit-cost accounting. At the local level, while input output modelling allows analysts to empirically estimate primary and secondary effects of salmon recovery actions, it is important to recognize that these are not valid measures of regional social welfare. Most certainly, these estimated effects do not persist in perpetuity.

One way of thinking about these issues is in terms of the paradigms of partial versus general equilibrium analysis. We have a tendency to think initially in terms of partial analysis -- the impact of lost waterway grain transportation on farmers, the impact of harvest restrictions on sport and commercial fisheries. In fact our analytic tool of choice, input-output analysis, is a partial equilibrium tool. I-O can be used to trace the backward linked effects on the rest of the economy of actions like salmon recovery. However I-O misses the general equilibrium aspects of recovery, because it implicitly assumes that demand for outputs and supply of inputs are infinitely elastic. A regional economy is a lot more flexible, re-employing idled resources, and substituting for lost outputs, than the partial equilibrium I-O model allows for.

REPLACEMENT COST VERSUS CONSUMERS SURPLUS

I want to illustrate this point with two further examples of the problems that partial equilibrium thinking can cause. I remember sitting in one of the interminable meetings of Dan Huppert's Economics Technical Committee. I was half listening to a discussion of the impact which salmon recovery would have on hydropower generation. Someone was making the point that such and such quantity of hydropower kilowatt-hours would be lost, and would have to be replaced by thermal power, at such and such "replacement cost". Along about then the internalized economist in me started to protest "but isn't there a demand curve for electricity? If hydropower lost to salmon recovery must be replaced by more expensive thermal power, won't this raise prices, and reduce demand, so that not all lost generation will need to be replaced?"

When I put up my hand and asked these questions, I think the initial reaction of the committee was to humor me. The partial equilibrium paradigm of using replacement cost is deeply entrenched in the minds of power planners. With the help of Ken Corum of the NW Power Planning Council, we eventually worked out the concepts of demand, pricing, and consumers surplus in an average cost pricing industry like electric power , and eventually the rest of the committee pretty much went along. Dan Huppert's report to NMFS (based on input from the Economics Technical Committee) did try to take a larger general equilibrium perspective on the power cost issue. We accomplished a paradigm shift on this issue.

BARGE TRANSPORTATION OF GRAIN

Another place where I became concerned about the effects of partial equilibrium thinking was in looking at the effects of river drawdown on barge transportation of grain. Dropping the levels of the lower Snake reservoirs for three to six months is advocated by some who argue that it would speed the passage of juvenile salmon downstream -- but this would also disable barge transportation of grain for the duration of the drawdown.

The initial reactions from partisans of barge transportation painted a horrendous picture of the impacts of drawdown. They argued that the same grain would have to move between the same points at the same times as without drawdown. If barges weren't available, then we would have to use rail and trucks to move the grain at sharply higher costs. They raised the spectre that farmers couldn't pay these rates and remain in business, so grain land in Idaho, Oregon and Washington would be idled.

I have argued that the farm economy has a lot more general equilibrium flexibility than that (see Hamilton, Casavant and Martin). If farmers were faced with a 30 cent a bushel or more difference in grain value net of transport price to Portland, I would expect the farmers to try very hard to move as much grain as possible during the months when the waterway is open. I would expect the ports to adjust in time to this new shipping pattern.

It is very unlikely that any land would exit from grain production as a result of river drawdown. There will be some costs passed on to farmers, which in a general equilibrium setting will show up as lower land prices, not as permanent unemployment of the land resource. My bottom line is that I think the true costs of river drawdown are far less than the estimates of some river transportation partisans -- largely because of the general equilibrium adjustments by all affected individuals.

CONCLUSIONS

This paper has addressed several conceptual issues which I have encountered as I have tried to understand the economics of salmon recovery. These have included the need to try to rigorously apply the with-without analytic framework, the need to look beyond impacts to a broader focus on costs and benefits, and more generally, the need to view things from a general rather than a partial equilibrium analytic framework. Confusion over these issues has clouded a lot of the economic analysis of salmon recovery. We need to understand and deal with these issues in our analysis if our work is to be theoretically defensible and truly relevant to the salmon recovery planning process.

REFERENCES

1. Hamilton, J.R., M. Martin, and K. Casavant, "The Effects of Lower River Reservoir Drawdown on Barge Transportation: Some Observations," Pacific Northwest Extension Bulletin #406, August 1992.

2. Hamilton, J.R., N.K. Whittlesey, M.H. Robison, and J. Ellis, "Economic Impacts, Value Added, and Benefits in Regional Project Analysis," American Journal of Agricultural Economics, May 1991, pp 334-344.

3. Hughes, D.W., and D.W. Holland, "Economic Impacts, Value Added, and Benefits in Regional Project Analysis: Comment" American Journal of Agricultural Economics, August 1993, pp 759-762.

4. Hamilton, J.R., N.K. Whittlesey, M.H. Robison, and J. Ellis, "Economic Impacts, Value Added, and Benefits in Regional Project Analysis: Reply," American Journal of Agricultural Economics, August 1993, p 763.

5. Huppert, D.D., D.L. Fluharty, and E.S. Kenney, "Economic Effects of Management Measures within the Range of Potential Critical Habitat for Snake River Endangered and Threatened Salmon Species," University of Washington School of Marine Fisheries, June 1992.

CONTINGENT VALUATION OF RARE AND ENDANGERED SPECIES: AN ASSESSMENT

T. H. Stevens, Martha K. Field, Thomas A. More
and Ronald J. Glass[1]

Introduction

Contingent valuation is playing an increasingly important role in decision making about rare and endangered species, but there are several problems with this approach.[2] Since most respondents are not very familiar with the commodity being valued, WTP depends in part on the information which is (or is not) provided in the survey. Response rates are often relatively low, and because of ethical concerns, some respondents may be either unwilling or unable to make meaningful tradeoffs between money and wildlife. Others may simply respond in terms of how they think the world "ought" to be (Opaluch and Grigalunas, 1992).

We begin with a brief overview and assessment of these issues. An alternative to the traditional CV method, conjoint analysis, is then examined.

Background

Contingent values for several rare or endangered species are shown in Table 1 (Brown, 1992)[3]. These values seem consistent with prior expectations; they vary by species and in general well known species appear more valuable than those which are not. However, many individuals have expressed concern about the validity of CV estimates for rare and endangered species (Sagoff, 1988; Harris et. al., 1989: Edwards, 1986).

---

[1] This research was sponsored, in part, by the Forest Service, U.S. Department of Agriculture.

[2] For example, pressure is mounting to incorporate provisions for benefit-cost analysis in the Endangered Species Act and CV is used for measuring damages and liability associated with oil spills and hazardous wastes.

[3] Brown (1992) found that 35% of the CV studies of nonuse values published in the U.S. since 1980 involved rare or endangered species.

One concern is that most respondents are probably not used to thinking in terms of making tradeoffs between money and endangered species.[4] Consequently, contingent valuation may function as "a kind of tutorial, building the monetary value as it elicits it" (Gregory, et. al., 1991, p.4). When viewed from this perspective, CV is seen as "an active process of value construction rather than as a neutral process of value discovery." (Gregory, et.al., p.5). Results from the CVM may therefore be very sensitive to the way in which choices are presented, survey format, and the information provided to respondents.

Another issue is that contingent value questions about rare or endangered species may induce decision making behavior dominated by ambivalence. Harris (1989) reminds us that wildlife are often viewed as, "either priceless or beyond market like transactions.." (p. 222), and Bishop (1985) notes that.."endangered species decisions must necessarily confront an important issue of intergenerational equity" (p. 209).

Contingent value questions about rare or endangered species may therefore produce conflict and ambivalence leading to avoidance, high rates of protest zero bids, and lexicographic behavior whereby tradeoffs between money and wildlife are not defined (Opaluch and Seagerson, 1989; Harper, 1989; Edwards, 1986 Stevens, et. al., 1991)[5].

A more troublesome problem is that respondents may not necessarily value the resource itself. Kopp (1992), for example, suggests that some individuals may be paying primarily to maintain ethical consistency.

Assume, for example, a simple individual utility function:

(1) $U = u(S,Y)$

---

[4] Gregory, et. al. (1991) argue that the well known embedding problem occurs because respondents do not have well defined monetary values for goods, such as wilderness and endangered species which are not priced in the market.

[5] The average response rate for the studies cited in Table 1 which sampled the general population was 35%. This is somewhat below average; Mitchell and Carson (1989) suggest that CV response rates average about 48%.

Where S is a binary variable representing existence of an endangered species, (S=1 if the species exists, 0 otherwise), and Y a vector of all other goods and services (income). Individual well-being or welfare, on the other hand, might be given by:

(2) $W = F(E) + U(S,Y)$

Where E is a vector of ethics. In particular, E=1 when behavior is consistent with this individual's code of ethics and E=0 otherwise.[6]

Following Kopp, assume that this individual is asked to vote for a program which would extract \$N to ensure preservation of an endangered species. Assume that $u(1, Y-N) < u(0, Y)$; this program produces a utility loss and from the perspective of neoclassical utility theory it would therefore be opposed. However, suppose that voting against this program is ethically inconsistent. This individual is then faced with a "no-win" situation; she must choose between the lessor of evils, and although a utility loss is involved, she may nevertheless be willing to pay to maintain ethical consistency. But, this individual's willingness-to-pay should NOT be interpreted as the economic value of preservation. Rather, it is the economic value of ethical consistency.

A somewhat different formulation is presented by Hollander (1990). Consider an individual with income Y who contributes amount N for the provision of a public good, G. The remaining income is allocated to a set of private goods, X.

(3) $X = Y - N \quad 0 \leq N \leq Y$

Following Hollander (1990), the supply of the collective good, G, is a function of the aggregate contribution from n individuals, but when n is large, each individual's contribution has a negligible effect on the amount of the collective good produced.

Assume that this individual's utility depends, in part, on sentiments, such as "social approval" associated with making the contribution, N. The individual's preference structure might then be given by:

---

[6] This structure is consistent with Sen's (1979) concept of commitment which drives a wedge between personal choice (utility) and personal welfare, W.

$$(4) \qquad U = U_x(X) + U_g(G) + U_s(s(N))$$

Where s(N) represents social approval. The value of utility when amount N is contributed is therefore:

$$(5) \qquad U_d = U_x(Y\text{-}N) + U_g(G) + U_s(s(N))$$

The value of utility when the donation is not made is given by:

$$(6) \qquad U_{nd} = U_x(Y) + U_g(G)$$

This individual is assumed to contribute amount N if, and only if:

$$(7) \qquad U_d \geq U_{nd}$$

The willingness-to-pay probability can then be expressed as: $(8) \qquad P_r = F(dV)$

Where F is the probability function for the random component of utility and dV is the utility difference:

$$(9) \qquad dV = U_x(-N) + U_s(s(N)) \qquad dV \geq 0$$

Willingness to pay (WTP) is therefore a function of the amount contributed, N, and social approval. The important feature emerging from this formulation is that utility difference, and hence the WTP probability, is <u>independent</u> of the collective good, G. Provision of the public good does not depend on this individual's contribution, and consequently, her WTP decision depends entirely upon utility derived from income relative to utility obtained from social approval associated with the act of contribution.

Similar formulations are presented by Andreoni (1990), by Kahneman and Knetsch (1992) and by Opaluch and Grigalunas (1992) who suggest that individual contributions for the provision of public goods may be motivated, at least in part, by desire for the "warm glow" associated with giving. Andreoni argues that some individuals may be motivated by "warm glow" only, others may value the public good, but care nothing for warm glow, and some may derive utility from both sources. The latter situation is an example of "impure altruism", and the resulting value estimate does not necessarily indicate the economic surplus derived from the good, G. As summarized by Opaluch and Grigalunas (1992);

> "Although impure altruism measures a form of personal benefit, it actually reflects the value of <u>doing good</u>, not the value derived from the good. Hence without taking care to identify the underlying motivation for responses, a CV survey can easily misinterpret the satisfaction obtained from contributing to a good cause as benefits associated with the specific commodity being described." (1992, p 4).

Empirical evidence about the meaning and interpretation of contingent values for endangered species is presented by Stevens, et al. (1994) who recently examined the motives underlying CV responses for bald eagle restoration in New England. When asked about the factors considered in deciding how much they would pay for bald eagle restoration, 31% of respondents cited household income and other financial commitments as most important. However, 17% cited environmental quality in general and 13% said that "doing my fair share" was most important. Moreover, many respondents considered more than one factor; 35% of those who cited income and financial constraints as most important also considered payment of fair share. Bids for preservation of rare or endangered species must therefore be interpreted carefully because contingent values appear to reflect the aggregate value of a bundle of attributes associated with paying for wildlife preservation. In other words, WTP includes payment for joint products, such as fairness, social approval, and ethical consistency, not necessarily associated with the resource itself.

Additional evidence of this phenomena is reported by Schkade and Payne (1994) who used verbal protocol analysis to examine the thought process of individuals when asked about their willingness-to-pay to protect migratory waterfowl from waste oil holding ponds.[7] Seventeen percent of respondents viewed their WTP as contribution to charity, 23% were concerned about environmental quality in general and 20% simply made up a number. Schkade and Payne conclude that "the WTP responses we observed seem to be constructed from a variety of considerations, including an obligation to pay a fair share of the cost of the solution and signaling concern for a larger set of environmental issues" (p. 88). Our overall assessment is that contingent values for rare and endangered species may be suspect. Most individuals are not used to making tradeoffs between money and endangered species. Ambivalence associated with ethical considerations may result in high rates of protest and non response. Moreover, it is very difficult to determine what people are actually paying for. Given these problems, examination of alternatives to the traditional CV seems appropriate. One alternative, to which we now turn, is conjoint analysis.

---

[7] Verbal protocol is a type of "think aloud" analysis used in psychological research.

Conjoint Analysis

Conjoint analysis, which is a modification of the referendum CVM, asks survey respondents to rate alternative programs. Since each alternative consists of a bundle of attributes, including price, marginal rates of substitution between attributes as well as the economic value of individual program attributes can be measured.

Following McKenzie (1990), assume an endangered species preservation program, Z, with M attributes:

(10) $Z = z(Z_i ,, \quad , Z_m)$

Where Z_i is the quantity of the ith attribute. In conjoint analysis product price or cost, P_z, is treated as just another attribute so that:

(11) $Z = z (Z_i ,, \quad , P_z)$

If the utility function is separable then:

(12) $U = U (Z_i , \quad - P_z + Y)$

Where Y is all other goods (income).

In this approach individuals are asked to rate alternative programs. For example, assume two programs, B and C, which differ only in terms of price and attribute Z_i. The utility difference between them is given by:

(13) $(U^b - U^c) = [U^b(Z_i^b+Y-P_z^b)]-[U^c(Z_i^c+Y-P_z^c)]$

If $U^b > U^c$, then program B is rated above C. Rating can therefore be expressed empirically as:

(14) $\text{Rating} = a(P_z) + B(Z_i) + e$

Where (14) represents the empirical utility function (McKenzie, 1993). Since the two attributes in (14) can be varied while leaving the rating (utility) constant, -B/a is the marginal willingness-to-pay for attribute Z_i (McKenzie, 1990; Magat, et al., 1988)

Previous empirical research suggests significant differences between conjoint and traditional CVM results. Ready, Whitehead and Blomquist (1991) used a polychotomous choice, PC, format in which respondents were given six responses to choose from, "definitely yes, probably yes, maybe yes, maybe no, probably no, and definitely no". When compared to the conventional format, PC resulted in slightly higher response rates and much higher estimates of WTP for two different amenities (preservation of wetlands and horsefarms).

Magat, Viscusi and Huber (1988) used both an open ended CV and a paired comparison format to value morbidity risk reductions associated with household chemicals (bleach and drain openers). In four applications, the paired comparisons approach produced higher valuations. Similar results were reported by Desvousges, et al. (1983) for water quality improvement in the Monongahala River.

Brown (1984) found that when asked for WTP subjects would pay more for ordinary commodities (cameras, etc.) than for environmental amenities (air quality and scenic quality). However, when asked to choose, most respondents rated amenities higher than commodities. More recently, Irwin, et al. (1993) found that WTP questions lead to relatively greater preference for improved commodities, such as TV's and VCR's, while choice questions yielded relatively greater preference for air quality.

Several factors may be responsible for these results. Compatibility and prominence effects are thought to be major causes of the preference reversals reported by Brown and by Irwin et al. (1993).[8] Irwin, et al. (1993) argue that,

> " The compatibility effect implies that when dollars are an available (recognizable) attribute of an object, they carry more weight or influence in determining an equivalent response that is also in dollars (e.g., cash equivalent, selling price) then they do in determining a response that is not in dollars (e.g., a rating of value or a choice). The prominence effect causes choice responses to be more dominated by prominent attributes than are pricing responses. This arises from the fact that choices are driven by reason and arguments to a greater extent than are pricing responses" (p. 6, 7).

Other factors may also be involved. As noted by McKenzie (1990), an important difference between conjoint and CV is that conjoint treats price as one of several attributes. Conjoint may therefore be particularly appropriate for valuing environmental amenities which people do not normally think of in monetary terms because it does not ask respondents to value the amenity <u>directly</u>. In the

[8] Preference reversals occur when CVM results yield a different ordering of options then when respondents are asked to choose among options.

context of endangered species valuation, conjoint does not require explicit tradeoffs between wildlife and money. Rather, these tradeoffs are implicit in the choices (ratings) made by respondents. Consequently, incentives for lexicographic decision making may be reduced. Conjoint analysis also allows respondents to express ambivalence. As noted by McKenzie (1993), "Conventional referendum CVM studies generally discourage respondents from indicating indifference or ambivalence, and estimation biases are likely to arise when indifference or ambivalence is expressed as nonresponse instead" (p. 593). Moreover, many individuals may be more familiar with making choices in a conjoint format. McKenzie (1990) argues that,

> "Respondents are generally more comfortable providing qualitative rankings or ratings of attribute bundles which include prices, rather than dollar valuations of the same bundles without prices. In treating price as simply another attribute, the conjoint approach minimizes many of the biases that can arise in open-ended contingent valuation studies when respondents are presented with the unfamiliar, and often unrealistic, task of putting prices on nonmarket goods" (p 112).

Conjoint analysis may therefore have several potential advantages for valuation of endangered species. However, we are not aware of any empirical CV/conjoint comparisons for rare or endangered species. A case study of Atlantic salmon restoration in New England was used to examine these issues.

Case Study

A survey was administered to 82 undergraduate students at the University of Massachusetts during the fall semester, 1993. Each individual received two CV questions about restoring Atlantic salmon to the Connecticut river. The first asked about WTP for a program which would restore Atlantic salmon throughout the lower half of the river. The second asked about WTP for a program described as being the same in all respects except that salmon would be restored throughout 90% of the Connecticut River system. These contingent valuation questions were followed by a question which asked respondents to rate five different restoration programs, (two of which were the same as those in the CV portion of the survey), on a scale of 0 to 9 with 9 representing programs the individual would definitely contribute to and 0 representing programs the respondent would definitely not contribute to.

All programs were described in terms of several attributes; cost (per person), region of river restored (% of river), expected salmon returns per year, probability of seeing salmon (low in all cases), and fishing opportunities (none in all cases). A copy of the survey is presented in the Appendix.

Seventy-six of the 82 students completed the questionnaire. CV results are presented in Table 2. The average WTP was $19.43 to restore salmon throughout the lower half of the river, and $27.89 to restore 90% of the river. The value of the 50 to 90% restoration increment is therefore $8.46.[9] Since each respondent also rated five restoration programs, 380 observations were available for conjoint analysis. The ordered probit procedure was used in which the dependent variable (ranking, 0-9) was regressed against program cost (cost), percent of river restored, salmon returns, 8 dummy variables accounting for rating intervals, MU1-MU8, and the mean of each respondent's five ratings (average rating).[10]

The results of estimation presented in Table 3 show that as expected, higher ratings are associated with larger salmon returns, a greater percentage of the river restored, and lower program costs. Marginal valuations for percent of river restored and per fish are given in column 3 of Table 3. A binary logit model was also estimated in which the dependent variable equals 1 for programs, if any, individuals would definitely contribute to (rating=9), and 0 otherwise (see Table 4). As expected, marginal valuations derived from this approach were quite similar to those obtained from the ordered probit analysis.

Value estimates for restoration of 50% and 90% of the river derived from the CV and conjoint analyses are compared in Table 5. The rating (conjoint) format produced <u>average</u> values for the 50% to 90% restoration <u>increment</u> which were 19 to 26 percent greater than those derived from the traditional CV; a finding which is generally consistent with previous research. However, 95% confidence intervals

---

[9] No protest zero bids or outliers were detected.

[10] The average rating variable facilitates comparison of ratings between respondents (see McKenzie, 1993).

for conjoint and traditional CV values overlap. Following McKenzie (1993) the ordered probit confidence interval was calculated from the inequality:

$$[b_i - b_c WTP]/\{ S_{2i} - S_i S_c WTP + S_{2c} WTP^2 \}^{.5} > t$$

Where b_i is the estimated coefficient for percent of river restored, b_c is the estimated restoration program cost coefficient, and S_i, S_c, $S_i S_c$ are coefficient variances and covariances (see McKenzie, 1993). Upper and lower bounds for conjoint values were $24.80 and $.80 (for the 50% to 90% restoration increment), while the corresponding CV confidence interval was $12.35 to $4.50.

From this perspective, CV and conjoint results are not different. But, when compared to CV, average values obtained from the rating format (logit estimates) for each individual restoration alternative (ie., 50% and 90% restoration) were less than the corresponding CV values, and several types of inconsistencies were observed when each individual's CV response is compared to that obtained from the rating format.

Forty-three percent gave a rating of less than 9 for identical programs they said they would contribute to in the CV portion of this study. Fifty-three percent preferred 90% restoration to 50% when asked in the CV format; that is, 53% were WTP more for 90% than for 50%. However, only 32% rated the 90% program higher than the 50% one in the rating format. And when each respondent's CV and rating responses are compared, 37% exhibited preference reversals.[11]

There may be several reasons for these differences. First, order effects could be a factor. The rating format followed the CV question and this may have effected ratings for individual programs. Also, respondents are presented a wider range of choice in the rating format. When responding to CV, 46% would pay $25 or more for the 50% restoration program, and 22% would pay $40 or more for the 90% program. But, when asked to rate alternatives, only 16% would definitely pay $25 for the 50% program and 11% would pay $40 for the 90% alternative. However, 42% would pay $40 for a third alternative which would restore salmon throughout the entire river (100% restoration). In essence the

[11] Since the rating question was expressed in dollar terms, the compatibility and prominence effects noted in previous studies of preference reversal ought to be minimal.

presence of a third, fourth, and fifth alternative in the rating format may have influenced the value of the first two.

Another factor is that the rating approach may provide less incentive for payment to maintain moral consistency, social approval, or 'warm glow'. This is because respondents are not forced to vote for or against programs in conjoint. Rather, rating allows respondents to express ambivalence by giving a rating between the endpoints (0 and 9). It is important to note that when asked in the rating format, 63% of respondents were not sure whether they would contribute.

Summary and Conclusions

Contingent valuation of rare or endangered species has been very controversial. Because ethical values are involved, CV may produce conflict resulting in nonresponse, payment to maintain moral consistency and related problems. Conjoint analysis may produce better results because compared to the traditional CV, tradeoffs between money and wildlife are implicit. Conjoint also allows for ambivalence and many people may be more familiar with making decisions in this format.

In a case study of Atlantic salmon restoration in New England, average values derived from conjoint were not statistically different from contingent values, but several important differences between individual's rating and CV responses were observed. This was not unexpected--as noted by Schkade and Payne (1994), "the sensitivity of CV results to various methodology factors (including discrepancies between WTP and choice) are completely consistent with over 20 years of research on the psychology of decision making" (p. 104).

We believe that conjoint is preferred from a conceptual perspective, but much more empirical evidence comparing conjoint and traditional CV formats is obviously needed.

Table 1. Summary of Recent CV Studies of Rare or Endangered Species

Author (Year)	Species	Mean Annual WTP, $, per household/year	
		Non-use value	Use-Value
Boyle and Bishop (1987)	Maintain and restore Bald Eagle Habitat in Wisconsin	$28.38[a] - $18.02	$46.93[a] - $57.29
Boyle and Biship (1987)	Striped Shiner	$5.55	NA
Brookshire, et al (1983)	Improvement of Grizzly Bear Habitat	$15.20	$5.80
Duffield (1992)	Wolf Recovery in Yellow-stone N P	$17.39	$5.48
Hageman (1985)	Avoid Reduction in Sea Otter Population	$13.62	$7.20
Hagen, et al (1991)	Protect Habitat of Northern Spotted Owl	$86.32	NA
King, et al (1986)	Survival of a Local Herd of Big Horn Sheep	$15.14	$2.00
Rubin, et al (1991)	Assure Existence of Spotted Owl	$49.72	NA

Table 1 (Continued)

Author (Year)	Species	Mean Annual WTP, $, per household/year Non-use value	Use-Value
Stevens, et al (1991)	Bald Eagle Restoration in New England	$15.81	$3.47
Stevens, et al (1991)	Atlantic Salmon Restoration in New England	$7.93	NA
Stoll and Johnson (1985)	Preserve Whooping Crane Habitat	$9.33[b] - $1.03	$7.54[b] - $9.64

Source: Thomas Brown. 1993. "Measuring Non Use Value: A Comparison of Recent Contingent Valuation Studies". W-133 Sixth Interium Report.

[a] Range depends on whether individuals would be able to use the resource.

[b] Range depends on whether individuals were visitors to wildlife refuge.

Table 2. CV Results for Salmon Retoration

Restoration Program	Average WTP ($)
50% of River (Lower Half)	$19.43
90% of River	$27.89
Marginal Value of 50% to 90% Increment	$ 8.46
n =	76

Table 3. Ordered Probit Conjoint Analysis of Salmon Restoration

Variable	Coefficient* (bi)		Marginal Valuation (bi/ b cost)
MU (1)	.17	(3.86)	
MU (2)	.37	(6.28)	
MU (3)	.69	(9.34)	
MU (4)	.93	(11.43)	
MU (5)	1.21	(13.69)	
MU (6)	1.49	(15.96)	
MU (7)	1.86	(18.99)	
MU (8)	2.23	(21.75)	
Average Rating	.359	(10.84)	
Cost	-.035	(3.33)	
% River Restored	.967	(2.08)	$27.63/100%
Salmon Returns	.00054	(4.42)	$ .015/fish
Constant	-.58	(3.13)	
Chi-squared	140.39	(4 df)	
n = 380			

*Absolute t values in parentheses

Table 4. Logit Model of Salmon Restoration

Variable	Coefficient* (bi)	Marginal Valuation
Average Rating	.44 (5.06)	
Cost	-.13 (4.36)	
% River Restored	3.34 (2.51)	$25.69/100%
Salmon Returns	.0011 (3.36)	$ .09/fish
Constant	-3.02 (5.30)	
Chi-squared	66.21 (4 df)	
n = 380		

*t Values in parentheses

Table 5. Comparison of Value Estimates: CV and Conjoint

Model	WTP ($) For Restoration Program 50% of River	90% of River	50-90% Increment
CV (Average WTP)	$19.43	$27.89	$8.46
Conjoint (Rating)			
1. Ordered Probit	-	-	$11.05
2. Logit*	$11.27	$21.54	$10.27

* Median (probability = .5)

REFERENCES

Andreoni, J. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal* 100:464-477.

Brown, T. 1993. "Measuring Non Use Value: A Comparison of Recent Contingent Valuation Studies" in *W-133 Benefits and Cost Transfer in Resource Planning, Sixth Interim Report*. Department of Agricultural and Applied Economics, University of Georgia.

Brown, T. 1984. "The Concept of Value in Resource Allocation." *Land Economics* 60:231-246.

Desuousges, W. H., et al. 1983. "A Comparison of Alternative Approaches for Estimating Recreation and Related Benefits of Water Quality." U.S. EPA, Washington, D.C.

Edwards, S. 1986. "Ethical Preferences and the Assessment of Existence Values: Does the Neoclassical Model Fit?" *Northeastern Journal of Agricultural and Resource Economics* 15(2):145-150.

Gregory, R., S. Lichtenstein, P. Slovic. 1992. "Valuing Environmental Resources: A Constructive Approach." Mimeo, Decision Research, Eugene, OR.

Harris, C. C., et al. 1989. "Improving the Contingent Valuation Technique: A Psychological Perspective." *Journal of Environmental Economics and Management* 17(3):213-229.

Harper, C. 1989. "Rational Roots of Irrational Behavior: Discussion." *Northeastern Journal of Agricultural and Resource Economics* 18(2):96-97.

Hollander, H. 1990. "A Social Exchange Approach to Voluntary Cooperation." *American Economic Review* 80(5):1157-1167.

Irwin, J. R., et al. 1993. "Preference Reversals and the Measurement of Environmental Values." *Journal of Risk and Uncertainty* 6:5-18.

Kahneman, Daniel and J. L. Knetsch. 1992a. "Valuing Public Goods: The Purchase of Moral Satisfaction." *Journal of Environmental Economics and Management* 22:57-70.

Kahneman, D. and J. L. Knetsch. 1992b. "Contingent Valuation and the Value of Public Goods: Reply." *Journal of Environmental Economics and Management* 22:90-94.

Kopp, R. 1992. "Ethical Motivations and Non Use Value." Mimeo.

Magat, W. A., et al. 1988. "Paired Comparison and Contingent Valuation Approaches to Morbidity Risk Valuation." *Journal of Environmental Economics and Management* 15:395-411.

McKenzie, J. 1993. "A Comparison of Contingent Preference Models." *American Journal of Agricultural Economics* 75:593-603.

McKenzie, J. 1990. "Conjoint Analyses of Deer Hunting." *Northeast Journal of Agricultural and Resource Economics* October:109-117.

Mitchell, R. C. and R. T. Carson. 1989. *Using Surveys To Value Public Goods: The Contingent Valuation Method*, Baltimore, Md:Johns Hopkins University Press.

Opaluch, James and Thomas Grigalunas. 1992. "Ethical Values and Personal Preferences as Determinants of Non-Use Values: Implications for Natural Resource Damage Assessments." Department of Resource Economics Staff Paper. University of Rhode Island.

Opaluch, J. and K. Segerson. 1989. "Rational Roots of Irrational Behavior: New Theories of Economic Decision Making." *Northeastern Journal of Agricultural and Resource Economics* 18(2):81-95.

Ready, R., J. Whitehead and G. Bloomquist. 1991. "Polychotomous Choice Valuation Questions" in *W-133 Benefits and Costs in Natural Resource Planning, Fourth Interim Report*, Department of Agricultural Economics, University of California, Davis.

Sagoff, M. 1988. "Some Problems with Environmental Economics." *Environmental Ethics* 10:54-74.

Schkade and 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.

Sen, A. 1979. "Personal Utilities and Public Judgements: Or What's Wrong with Welfare Economics?" *The Economic Journal* 89:537-558.

Stevens, Thomas H., et al. 1991. "Measuring the Existence Value of Wildlife: What Do CVM Estimates Really Show?" *Land Economics* 67(4):390-400.

Stevens, Thomas H., et al. 1994. "Interpretation And Temporal Stability of CV Bids for Wildlife Existence: A Panel Study." *Land Economics* (forthcoming August, 1994).