WESTERN REGIONAL RESEARCH PUBLICATION

W-133
Benefits and Costs Transfer in Natural Resource Planning

Tenth Interim Report

1997

Compiled by
Jeffrey E. Englin

Department of Applied Economics and Statistics
University of Nevada
Reno, NV 89557-0105

CONTENTS

Introduction

This volume contains the proceedings of the 1997 meetings of the W-133 Western Regional Research Project, "Benefits and Costs Transfer in Natural Resource Planning" held in Portland, Oregon. The goal of this project is to provide estimates of the benefits of the environment for inclusion in cost-benefit analysis of public policies. The results should inform policy makers and provide guidance to those who undertake such work in the future.

The meeting in Portland was attended by a wide range of land grant academic faculty, non-land grant academic faculty, federal decision-makers and analysts, and state and local analysts. The work presented in this volume represents on-going research at land grant institutions. It indicates the richness and quality of this area of research and clearly demonstrates its relevance to public policy in the United States.

Many people worked provide successful meeting. Billye French and Stephanie Fletcher provided the administrative support in the Department of Applied Economics and Statistics at the University of Nevada, Reno. The session chairs worked hard to assure an orderly meeting. I believe that they accomplished their goal, and I am grateful to them. I appreciate the help provided by Janet Lutz during the meeting. She made an important contribution to the organization of the meeting. Nicki Wieseke did the bulk of the work generating the proceedings. Finally, but most importantly, the authors provided a wide range of high quality work to include in the proceedings.

Jeffrey Englin
Department of Applied Economics and Statistics
University of Nevada
Reno, Nevada

THE AGGREGATION OF CONDITIONAL RECREATIONAL DEMAND SYSTEMS

J. Scott Shonkwiler and W. Douglass Shaw
Department of Applied Economics and Statistics/Mail Stop 204
University of Nevada
Reno, NV 89557

e-mail: jss@scs.unr.edu and wdshaw@unr.edu

* Our thanks go to Dan Hellerstein and Jeff LaFrance for their comments on a related manuscript and to Kerry Smith for his thoughts on issues relating to this paper. This research was partially supported by the Nevada Experiment Station via the W133 Regional Research Project.

Abstract

The random utitlity model (RUM) is commonly used to represent the individual's allocation of trips to a set of recreation areas. Empirical application of the RUM is performed using the conditional logit model. This model provides well-known measures of per-trip consumer surplus, but scaling up these measures to aggregate seasonal or annual values is problematic because the underlying multinomial logit model conditions on the total number of trips to all sites. As a result interest has focused on linking the conditional logit trip allocation model to a model of aggregate demand using a price index derived from the RUM. Using results on the logit representative consumer of Anderson et al., this paper shows that a utility theoretic aggregate price index that is consistent with a logit allocation model model does not exitst when the aggregate good is defined as total recreational trips. If aggregate demand is defined as total recreational travel and if the conditions for Hicks composite commodity theorem are satisfied, then it can be shown that trip allocation and total travel demand can be determined in a utitlity theoretic manner and welfare measures can be derived. The paper presents a conditional indirect utility function which links a site allocation model of logit form to a proper aggregate demand model.

THE AGGREGATION OF CONDITIONAL RECREATIONAL DEMAND SYSTEMS

1. INTRODUCTION

In this paper we explore the micro-theoretic linkage between the popular multinomial logit site allocation model (McFadden [12]) and a total trip demand model, all applied to a recreation context. Many recreation modelers have raised the point that consumer's surplus measures should come from some aggregate demand function rather than from the site-specific demands, because the former allows total seasonal consumption to change in response to site quality and price changes and the latter does not. Intuitively, when one only has per-trip welfare measures, some assumption must be made about whether and how these can be added together to arrive at a welfare measure that can be interpreted as an annual (seasonal) maximum willingness to pay (WTP) to bring about some change. By deriving the welfare measure from the aggregate function, we can avoid problems associated with these restrictive assumptions (see Morey [14]).

An ongoing question is how to model aggregate demand given that the data are detailed enough to provide information on site-specific demands. In this situation the data are rich enough to allow calculation of a travel cost to each individual recreation site, and it seems logical that this information should be exploited when developing a price for the aggregate model. Thus a number of analysts have tried to determine the appropriate aggregate price when site-specific demands have been modeled using the multinomial logit specification. Several of these price indexes are reviewed below.

We next consider the work of Anderson et al. [1] who begin with a representative consumer utility function and derive commodity demands (shares) which have the form of multinomial logit probabilities. The structure of this utility function is shown to preclude being able to derive a proper price index for the aggregate good, although a result of Gorman [9] is used to suggest a shadow price for the aggregate good.

These findings lead us to consider alternative ways to define aggregate recreation demand. Recognizing that the assumption that all price changes within the aggregate good are proportional is both reasonable and useful, we show how Hicks' aggregation theorem can be used to construct an aggregate measure used to link micro demand (trips to each site) with "macro" demand (total trips). It turns out that a nice argument can be made that total travel, rather than total seasonal trips, is the relevant aggregate good in such recreation modeling. We next develop a price for the aggregate trips model using a specific utility function which yields site specific shares of multinomial logit form and a corresponding aggregate recreation demand model. Both site choice and total recreation demand are derived from a single, integrated utility maximization problem. While this is generally applicable to many underlying structures, we focus on the conditional logit model for the site choices, as this is a utility-theoretic and popular specification.

2. BACKGROUND

A conventional recreation site choice model is the multinomial logit model of McFadden [12]. This model, while quite popular because of its attractive features in dealing with site allocation, yields a "per-trip" or per-choice occasion consumer's surplus measure. Some developers of recreation demand models attempt to empirically link site allocation and total seasonal trip demand in one "combined" model, largely in hopes of getting consumer's surplus from the aggregate demand function.[1] The idea is appealing, as failure to establish this linkage requires assumptions about how total trips for the season change (or more often how they are assumed to not change) when the price or characteristics at one or more sites change. For example, it is possible that an

[1] See for example Bockstael, Hanemann and Strand; Yen and Adamowicz; Hausman, Leonard and McFadden; Feather, Hellerstein and Tomasi; Parsons and Kealy; and the application by Shaw and Jakus.

"improved" site results in a smaller number of seasonal trips being taken, particularly if the goal is a form of consumptive recreation like fishing or hunting (see Englin, Lambert and Shaw [7]), because the individual gets higher utility from a single site visit than before the improvement and thus may not require as many total seasonal trips. Linked models are also appealing because the potential exists for establishing a clearer connection between specific site prices and some aggregate price.

Many of the combined or linked studies assume McFadden's [12] multinomial logit model to be appropriate for the site allocation portion because visitation data are discrete and the model can be easily used to estimate exact per-trip welfare measures for site quality changes (we ignore the additional and tangential issue of allowance for income effects here). For the total seasonal trips portion of the combined model a count data approach is frequently assumed, as the non-negative integer properties of it are desirable in the statistical treatment of the dependent variable(e.g. Hellerstein [11]). Such models beg the question: what is the correct price (denoted P) for all of the trips taken in a season?

To put the various proposed aggregate price indexes in context it is necessary to develop the multinomial logit site demand model. Assume that an individual's indirect utility for the j^{th} ($j=1, 2, \ldots, J$) site can be represented by

$$U(Y-p_j, z_j, \varepsilon_j) = V(Y-p_j, z_j) + \varepsilon_j \qquad (1)$$

where Y is an income measure for the individual, p_j is the individual's travel cost for a visit to the j^{th} site, z_j is a (vector) measure of characteristics associated with the j^{th} site, and ε_j is an unknown, idiosyncratic term associated with the individual. Under the linear in income case the systematic part of the indirect utility can be written $V = |\beta|(Y-p_j) + \gamma z_j$. McFadden [12] has shown when the ε's have a joint cumulative distribution of generalized extreme value form that the choice probability of visiting the j[th] site has the representation

$$\pi_j = \exp(\alpha_j - \beta p_j) / \sum_{i=1} \exp(\alpha_i - \beta p_i) \qquad \beta < 0, \ \alpha_i = [illegible] \qquad (2)$$

With the notation developed above, we can now present several price indexes proposed for the aggregate demand model in Table 1.

Table 1. A Comparison of Price Indexes Used in Modeling Aggregate Demand

STUDY	PRICE INDEX
Bockstael, Hanemann, and Strand [2]	$P_1 = \ln \sum_{i=1}^{J} \exp(\alpha_i - \beta p_i)$
Hausman, Leonard, and McFadden [10]	$P_2 = \beta^{-1} \ln \sum_{i=1}^{J} \exp(\beta p_i)$
Feather, Hellerstein, and Tomasi [8] and Parsons and Kealy [15]	$P_3 = \sum_{i=1} \pi_i p_i$

Some researchers (eg. first Bockstael, Hanemann, and Kling [2]; then Yen and Adamowicz [21], and recently Shaw and Jakus [17]) assume that the appropriate price index in the total seasonal trips demand function is the inclusive value (IV) from the conditional logit model which is labeled P_1 in Table 1. Alternatively, Hausman, Leonard and McFadden (HLM) [10] use the inclusive value (under the restriction that $\alpha_i = 1, \forall i$) scaled by the inverse of the price coefficient as the price in their aggregate trips demand function. P_2. This index has the interpretation as the negative of the multinomial consumer surplus per trip. Though HLM [10] state that their scaled inclusive value as a price index is consistent with two-stage budgeting for their combined models. they fail to specify a conditional indirect utility function that implies an expenditure function that is linear homogeneous in prices. Smith [19] has pointed out that a consequence of their specification is that it satisfies the conditions of two-stage budgeting only when a single site is considered.

The issue of the appropriate price index remains. despite proposals for other indexes such as Feather. Hellerstein and Tomasi [8]. or Parsons and Kealy [15] who both use the sum of the site prices weighted by the probabilities of visits to the j^{th} site for their aggregate price term. P_3. While this index appears to be the only one that is linear homogeneous in site-specific prices. this is not the case since the π_i's themselves depend on prices as shown in equation 2. While all theses indexes reflect thoughtful work. we believe that the issue of an appropriate aggregate price index clearly remains because not all of the above indexes can be correct. and because adequate justification for these indexes has not been presented.

What is perhaps of most interest in this paper is therefore the correct price in an aggregate trips demand function. A price for a group of goods. or price index. enters into many formal derivations for demand systems. For example. the geometric mean of the prices of all the goods in the group can be interpreted as a cost of living index in the linear expenditure system (see Deaton and Muellbauer [5] for example), and the sum of the log of the group prices weighted by their budget share is a general price index in Stone's model [20]. So. while the exact aggregate price may clearly differ depending on assumptions made. it flows from a formal (micro-theoretic) derivation in each case. We provide justification for an index below, paying particular attention to whether and how the conditional logit demand system can generate both an aggregate demand and aggregate price. As a preview, what may be a bit surprising to some is that the conventional logit derivation does not yield a particularly useful structure for the aggregate demands. We offer an alternative that is more useful.

3.0 MICROECONOMIC THEORY

Aggregate recreation demand has traditionally been defined as the individual's total number of visits to set of recreation sites.

$$Q = \sum_{i=1}^{J} X_i \qquad \text{where the } X_i \text{ represents the individual's demand for the } i^{th} \text{ site.} \qquad (3)$$

Alternatively aggregate recreation demand may be defined as

$$T = \sum_{i=1}^{J} \delta_i X_i \qquad \text{where } \delta_i \text{ represents the individual's distance to the } i^{th} \text{ site.} \qquad (4)$$

Note that (4) casts aggregate demand in terms of total recreational travel, rather than total recreational visits. While attention has focused on obtaining a suitable price index for (3), researchers have neglected the fact that summing trips may be a misleading enterprise because different sites may have substantially different prices. On the other hand (4) seems more compatible with the basic underpinnings of the travel cost model of recreation demand--namely that distances traveled to recreation sites are related to site prices, hence (4) is proportional to total recreational expenditures. To illustrate the consequences of defining aggregate recreation demand in terms of (3) as opposed to (4), we begin with an examination of a particular conditional utility function that yields a system of share equations for the individual sites that has the multinomial logit probability form . We can then explore the aggregate demand that stems from this conditional demand system and its corresponding price.

The Logit Representative Consumer

Following Anderson, De Palma and Thisse [1], the direct conditional utility function for the individual which is consistent with the choice probabilities in (2) is:

$$U = -\beta^{-1}\sum_{i=1}^{J}\alpha_i X_i - \beta^{-1}\sum_{i=1}^{J} X_i(\ln X_i - \ln Q) - X_0, \qquad \beta < 0,\ \alpha_i > 0\ \forall i \tag{5}$$

where X_0 is the demand for a composite good, and as defined previously X_i is the number of trips taken to site i. Let Y be total household income and Q be the total number of trips taken to all sites. If the amount Q is considered as exogenously given (Anderson et al.[1]), the Lagrangean for the consumer's optimization problem is

$$L = U + \lambda[Y - p_0 X_0 - \sum p_i X_i] - \gamma[Q - \sum X_i] \tag{6}$$

The price of X_i is p_i, and the price of X_0 is p_0. Solution to this problem yields the site demands. The first order condition for the j[th] good can be written $\ln X_j - \ln Q = \ln \pi_j = \alpha_j + \beta p_j p_0^{-1} + \beta\gamma - 1$ when the value $1/p_0$ is substituted for λ. It is convenient to express these in share form (see Morey [13]) as an early share example in the recreation literature). Exponentiating and summing the above yields a solution for γ which when substituted into the first order condition gives the demands in share form:

$$\pi_j = \exp(\alpha_j + \beta p_j p_0^{-1}) / \sum_{i=1}^{J}\exp(\alpha_i + \beta p_i p_0^{-1}) \tag{7}$$

Note that in a typical recreation analysis the p_o would be set to unity, yielding a familiar looking form for the probability (see equation 2).

To solve for the optimal Q, the Lagrangean in (6) can be optimized with respect to Q. But as pointed out by Anderson, de Palma and Thisse [1], this does not admit an interior solution for Q. Either all income is spent on the composite commodity or Q. The upshot is that there is no information from this system to determine how the site-specific prices need to enter the aggregate demand function. Thus there is no single optimizing specification that generates both logit choice probabilities for the individual sites and a solution for Q that explicitly depends on the individual p_j.

However a shadow price for Q can be obtained by deriving the conditional indirect utility function. V( ), from substitution of the optimizing $X_i = \pi_i Q$ into (5).

$$V = Y/p_0 - \beta Q \ln \sum \exp(\alpha_j - \beta p_j p_0^{-1}) \tag{8}$$

The shadow price for the aggregate demand from this indirect utility function can be derived following some results of Gorman [9] relating to conditional indirect utility functions. Gorman [9] defines the shadow price of the fixed good as:

$$-\frac{\partial V/\partial Q}{\partial V/\partial Y} = \beta^{-1} p_0 \ln \sum \exp(\alpha_j - \beta p_j p_0^{-1}) \tag{9}$$

This expression can be simplified if the price of the composite good can be set to unity. Doing so, we are left with an aggregate price of:

$$P = \beta^{-1} \ln \sum \exp(\alpha_j - \beta p_j) \tag{10}$$

Interestingly, this expression is the (negative) of the consumers surplus per trip that flows from the multinomial logit model. As the marginal utility of income is usually assumed constant. this β is the price coefficient. and so. with the exception of allowance for the site-specific constant. α_j, we are left with the same aggregate price as used by HLM [10]. The shadow price in (10) is proportional to that suggested by BHS [3], though the presence of the multiplier β^{-1} means that the price coefficient in their aggregate demand model is of opposite sign and likely overstated (if $-1<\beta<0$) by using P_1 rather than P.

Consistent Aggregation

We have suggested that total seasonal trips, obtained by summing up the trips to each site visited, characterize an aggregate good that cannot be associated with a price index (as a consequence of optimizing a single utility function) when site choice probabilities are of multinomial logit form. Conditions for existence of a well-defined aggregate commodity are provided in Diewert [6], who notes that the "use of aggregates in the theory of consumer demand can be justified, provided all price changes within an aggregate are proportional". The motivation for defining aggregate recreation demand in terms of total recreational travel (see equation 4) should now become clearer. When δ_i represents the individual's (round-trip) distance from the i^{th} site, then the price (travel cost) for the i^{th} site is proportional to δ_i given that we define $p_i = \rho\delta_i$. Here ρ is the marginal price per unit of distance traveled. In general it would be expected that ρ would not change for the various sites. In this manner, the travel cost model of recreation demand appears to provide sets of goods for which the composite commodity theorem of Hicks is naturally satisfied. Recognition of this feature is the key to consistent aggregation of site demands.

First consider the disaggregated utility maximization problem:

$$\text{Max } U(\mathbf{X}, X_0) \text{ subject to } Y = \mathbf{p}'\mathbf{X} + p_0X_0$$

where X and $\mathbf{p}$ are J element vectors of site visits and prices, respectively. Assume the utility function $U(\mathbf{X}, X_o)$ satisfies the minimum regularity conditions of Diewert [6], and denote the optimizing values of $\mathbf{X}$ by $\mathbf{X}^*$. Then the optimal value of T from this optimization problem can be written $T^* = \delta'\mathbf{X}^*$. Alternatively, we can consider what Diewert [6] calls the aggregated utility function

$$\text{Max } U(\mathbf{X}, X_0) \text{ subject to } T = \delta'\mathbf{X}$$

where δ is a fixed vector of constants. Notice that the variation in the price vector is constrained by the relationship $\mathbf{p} = \rho\delta$. By fixing T and X_0 and then maximizing this latter utility function with respect to $\mathbf{X}$, the optimal levels of $\mathbf{X}$, conditional on T are obtained. The conditional indirect utility function $V(T,X_0; \delta)$ is then obtained by substituting the expressions for the optimizing values of $\mathbf{X}$ into the utility function. Here, and perhaps slightly atypical for an indirect utility function, we note that V ( ) is conditioned on both T and X_0. Properties of V( ) are dependent on the regularity conditions for U( ): V is continuous, non-increasing and quasi-convex in δ, and homogenous of degree zero in T and δ. Note that conditional utility functions are discussed and conditional demands are developed by Pollak [16], among others. These conditional demands are, as Pollak [16] states, "directly relevant to the analysis of consumer behavior in the short run, when fixed commitments prevent instantaneous adjustment to the long run equilibrium..." (p. 60). The demands can conditioned on expenditures or on a fixed quantity, as in the case of rationed goods. If an individual's allotment of a preallocated good remains fixed, then a well-behaved conditional utility function can be specified with the preallocated good as one of the arguments.

An especially attractive feature of this Hicks/Diewert framework is that if we start with the conditional indirect utility function, say $V(T,X_0; \delta)$, and this is then maximized with respect to T and X_0, with the relevant constraint being $Y = p_0X_0 + \rho T$, then the resulting T^* is identical to that obtained from the disaggregated utility maximization problem above. Therefore, we could in fact begin with a conditional indirect utility function $V(T,X_0; \delta)$, assuming that it is nonincreasing and quasi-convex in δ, and homogenous of degree zero in δ and T, and do the following:

(i) Apply Roy's Identity to $V(T,X_0; \delta)$ to derive the conditional demands, $\mathbf{X}^*$

(ii) Maximize $V(T,X_0; \delta)$ subject to a budget constraint to obtain T^*, and

(iii) Invert the indirect utility function $V(T^*,X_0^*)$ to obtain the expenditure function which then can be used to do the desired welfare analysis.

In the next section, we apply this framework to the problem of obtaining the demand for the aggregate good given an allocation model that has the multinomial logit form.

4. AN EXAMPLE

We begin by specifying a conditional indirect utility function for an individual with the particular form

$$V(T,X_0;\delta) = \gamma T^{\beta} \sum_{i=1}^{J} \exp(\alpha_i - \beta \ln \delta_i) + (\theta + X_0)^{\beta} \tag{11}$$

where here $\beta > 0$, $\gamma > 0$, and $\theta > 0$. The parameter γ is a weight which determines the relative importance of recreation in this utility function, and as will be seen the parameter θ will affect the form of the demand for the aggregate good, T. Note that this specific conditional indirect utility function is non-increasing and quasi-convex in δ, and it is homogeneous of degree zero in T and δ. Using Roy's identity yields conditional demands of the form:

$$X_j^* = {T \exp((\alpha_j - \beta \ln \delta_j)} \Big/ {\delta_j \sum \exp(\alpha_i - \beta \ln \delta_i)} \tag{12}$$

or, in share form we obtain:

$$\pi_j = {\exp((\alpha_j - \beta \ln \delta_j)} \Big/ {\sum \exp(\alpha_i - \beta \ln \delta_i)} \tag{13}$$

Now, these equations are of the conditional logit form, and α_i and β can therefore be determined from the site allocation model conditioned on T.

To find the optimal amounts of T and X_0 we can maximize the conditional utility indirect function $V(T,X_0;\delta)$ with respect to T and X_0, subject to the usual constraint: $Y = p_0X_0 + \rho T$. The resulting T^* is:

$$T^{*} = \frac{\theta + Y / p_0}{\left[\rho^{-1} p_0 \gamma \sum \exp(\alpha_i - \beta \ln \delta_i)\right]^{1/(\beta-1)} - \rho / p_0} \tag{14}$$

The expenditure function for the above system can be obtained by substituting the optimizing expressions for T^* and X_0^*, where X_0^* is of the form $X_0^* = Y / p_0 - \rho T^*$, into the indirect utility function. There are several important implications of such a system.

First, a cross sectional study would typically allow normalization of p_0 to unity. Aggregate, as well as site-specific, demands could be estimated simultaneously, as they share the parameters α_i and β. Note also that γ and θ could be parameterized to depend on characteristics of the individual, such as age, gender, or education. The aggregate demands from this system will be homogenous of degree zero in prices and income. However, if $\theta = 0$, then the aggregate demands will be homothetic (otherwise, they are quasi-homothetic). From this system, we can use the expenditure function to obtain the exact measures of consumer's surplus, the compensating and equivalent variation (CV and EV, respectively). Of particular interest is the feature that these welfare measures are not conditional either on T or on the allocation of income at a higher stage of budgeting.

5. CONCLUSION

The literature now contains several possible indexes for use in an aggregate demand (trips) function, not all of which can be correctly interpreted in a consistent overall framework. Our derivation of an aggregate good that is consistent with the conditional logit model for recreation and its resulting aggregate price is based on travel (rather than aggregate trips) as the aggregate good. The accompanying aggregate price is slightly different than what has been seen before so one might question whether the welfare measures based on other aggregate prices are meaningful. Intuitively, our model is consistent with behavior that suggests an individual optimizes by thinking about the total amount of travel to be done over the course of the season. This assumption about recreating behavior is perhaps no more objectionable than the usual one.

REFERENCES

1. Anderson, S.P.; A. de Palma; J.-F. Thisse. 1988. A representative consumer theory of the logit model. *International Economic Review* 29: 461-66.

2. Bockstael, N.; W.M. Hanemann; C. Kling. 1987. Estimating the value of water quality improvements in a recreational demand framework. *Water Resources Research* 23: 951-60.

3. Bockstael, N.; W.M. Hanemann; I.E. Strand. 1984. "Measuring the benefits of water quality improvements using recreational demand models." Vol. II of Benefits Analysis Using Indirect or Imputed Market Methods. EPA Contract CR-811043-01-0.

4. Creel, M. and J. Loomis. 1990. Theoretical and empirical advantages of truncated count data estimators for analysis of deer hunting in California. *Amer. Journal of Agri. Econ.* 72: 434-41.

5. Deaton, A. and J. Muellbauer. 1980. Economics and Consumer Behavior. Cambridge: Cambridge Univ. Press.

6. Diewert, W.E. 1978. Hicks' aggregation theorem and the existence of a real value-added function. Chapter III.2 in Production Economics: A Dual Approach to Theory and Applications (M. Fuss and D. McFadden (eds.)). Amsterdam: North Holland Publishing Co.

7. Englin, J.; D. Lambert; W.D. Shaw. 1997. A structural equations approach to modeling consumptive recreation demand. Forthcoming, *Journal of Environ. Econ. and Manage.*

8. Feather, P.; D. Hellerstein; T. Tomasi. 1995. A discrete-count model of recreational demand. *Journal of Environ. Econ. and Manage.* 29: 214-227.

9. Gorman, W.M. 1976. Tricks with utility functions. Chapter 11 in Essays in Economic Analysis, M.J. Artis and A. R. Nobay(eds). Cambridge: Cambridge Univ. Press.

10. Hausman, J.; G. Leonard and D. McFadden. 1995. A utility-consistent, combined discrete choice and count data model: assessing recreational use losses due to natural resource damage. *Journal of Public Economics* **56**: 1-30.

11. Hellerstein, D.M. 1991. Using count data models in travel cost analysis with aggregate data. *Amer. Journal of Agri. Econ.* (August): 860-866.

12. McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. A chapter in Frontiers in Econometrics, P. Zarembka (ed.). New York: Academic Press.

13. Morey, E.R. 1985. Characteristics, consumer surplus and new activities: a proposed ski area. *Journal of Public Economics* **26**: 221-36.

14. Morey, E.R. 1994. 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. *J. of Environ. Econ. and Manage.* 27: 257-70.

15. Parsons, G.R. and M. Kealy. 1995. A demand theory for number of trips in a random utility model of recreation. *Journal of Environ. Econ. and Manage.* 29: 357-367.

16. Pollak, R.A. 1969. Conditional demand functions and consumption theory. *Quarterly J. of Economics* 83: 60-78.

17. Shaw, W.D. and P. Jakus. 1996. Travel cost models of the demand for rock climbing. *Agricultural and Resource Econ. Review* (October): 133–42.

18. Smith, V.K. and R. Von Haefen. 1997. Welfare measurement and representative consumer theory. Discussion paper. Resources for the Future.

19. Smith, V.K. 1997. Combining Discrete Choice and Count Data Models: A Comment. Working Paper, Duke University.

20. Stone, J.R.N. 1954. The measurement of consumers' expenditure and behavior in the United Kingdom, 1920-1938, Vol. I., Cambridge: Cambridge Univ. Press.

21. Yen, J.S. and V. Adamowicz. 1994. "Participation, Trip Frequency and Site Choice: A Multinomial Poisson Hurdle Model of Recreation Demand." *Canadian Journal of Agricultural Economics* 42: 65-76.

Combined Revealed and Stated Preference Analysis Involving Discovery of New Attributes: The Case of North American Aboriginal Artifacts.

Peter C. Boxall
Canadian Forest Service,
Natural Resources Canada

Jeffrey Englin
Department of Applied Economics and Statistics
University of Nevada-Reno

and

Wiktor L. Adamowicz
Department of Rural Economy
University of Alberta

20 February 1998

The authors thank the participants at the 25th Workshop of the W-133 Project held in Portland, Oregon in March 1997 for their comments on this manuscript and gratefully acknowledge David Watson for creative field work in the Shield country and for assisting with the survey design and administration. Dan Benoit, Kent Baylis, Jennifer Lidgett and Janet Lutz also provided valuable field assistance. This research was funded by the Canada-Manitoba Partnership Agreement in Forestry and the Socioeconomic Research Network of the Canadian Forest Service.

Combined Revealed and Stated Preference Analysis Involving Discovery of New Attributes: The Case of North American Aboriginal Artifacts.

This paper examines the hypothetical discovery of attributes in recreation site choice models using a joint revealed-stated preference data collection and modelling process. The empirical application involved the potential discovery of aboriginal rock paintings along wilderness canoe routes in the eastern Manitoba. A 4 year study of wilderness recreation trips provided the opportunity to carefully design a stated preference experiment in which canoeists were asked if they would change their site choices in response to the presence of two types of rock paintings: a "pristine" painting and another spoiled by human vandals. The resulting stated site preferences (with the new attributes) were combined with the revealed site preferences (without the attributes) in the econometric analysis. The results suggest that preferences over the SP and RP models were not statistically different. Welfare measures for the presence of "pristine" paintings range from $4.79 - $6.81 per trip, and are about 12-13 times greater than those for vandalised paintings.

Introduction

A challenge in the valuation of environmental amenities is the *ex ante* measurement of values associated with undiscovered goods and services. The behaviour inherent in any revealed preference information is, of course, associated with the *ex ante* situation. Newly discovered goods and services such as new species or cultural artifacts will result in new alternatives and/or new attributes at existing alternatives that may affect future behaviour. Since these goods or services are currently unknown or unavailable, there is no revealed preference information to use in their valuation. As a result, one is forced to rely upon stated preference information to examine the new attribute or alternative. Nevertheless, it is important to maintain consistency with revealed behaviour associated with the *ex ante* situation. The challenge is to acknowledge and exploit all of the information available when valuing newly discovered attributes.

A variety of modelling frameworks have been proposed to analyse combined revealed and stated preference data (e.g. Cameron (1992); Englin and Cameron (1996); Adamowicz et al. 1994, 1997). The Admowicz et al. (1994, 1997) random utility model (RUM) framework is especially appealing in settings where most individuals make a single trip. In this setting the single trip nature of pure random utility models is less troubling than in other contexts (see Morey (1994)). The appeal of the RUM is its ability to handle substitution between site attributes and the direct measurement of economic welfare, while retaining the ability to econometrically test the consistency of the revealed and stated preference components of the model.

This analysis examines the potential discovery of aboriginal rock paintings along wilderness canoe routes in the Canadian Shield. Anthropological scholars call these paintings pictographs because they represent picture writing, not works of art. These drawings were used to communicate among individuals or with the spirit world by the Algonkian nation. About 400 pictographs have been found on rock faces along water courses in the Canadian Shield. Anthropologists believe some of these pictographs to be 2000 years old (Rajnovich 1994). Archeological discoveries of pictographs are still occurring, and new pictographs continue to be catalogued periodically. While pictographs have spiritual and cultural significance to Aboriginal peoples, they are also highly sought by wilderness recreationists who consider them an important feature of a wilderness experience in the Canadian Shield (Boxall unpublished). The tension between value of the pictographs to users who experience them and the increased risk of vandalism to those pictographs is an increasing management concern.

The paper proceeds by developing the RUM used in this analysis in the next section. This model directly incorporates revealed preference information about existing recreation site attributes and stated preference information about, as yet, undiscovered site attributes. In this section the combined revealed–stated preference approach is developed. This is followed by a description of the data used in this analysis. The empirical section applies the combined revealed preference-stated preference model to the hypothetical discovery of aboriginal rock paintings along water courses in a wilderness area. The empirical application also examines the effect of vandalism on the benefits generated by the rock paintings.

Theory

Consider a recreationist who makes a choice from a set of C possible sites. The probability (□) that site *j* will be visited is equal to the probability that the utility gained from visiting *j* is greater than or equal to the utilities of choosing any other site in C. In this framework indirect utility consists of the sum of two components: an observed component, V_j, and a random component e_j. The probability of selecting site j can be written as :

$$\pi(j) = \Pr\{V_j + e_j \geq V_k + e_k \forall \text{ k in C}\}.$$

Empirical implementation of (1) requires the selection of a distribution to characterise the random component of the model (e_l). The conditional logit model can be used to estimate these probabilities if the random components of the indirect utility functions are assumed to be independently distributed with a Type-I Extreme Value distribution (Weibull).

This model is typically estimated with the observable component, V_j, expressed as a linear function of m site attributes and the cost of visiting a site. A new attribute introduced into this framework will take the form of an additional attribute in the indirect utility function:

$$V_j = \sum_1^m \beta_m X_{jm} + \alpha X_{jnew} + \gamma(Y_n - p_j),$$

where X_{jm} represents existing choice based attributes, $X_{j\,new}$ represents a new attribute, Y_n is income, p_j is the cost of visiting site j, and β, α, and γ are unknown parameters. However, by definition, the only revealed preference data available is based on behaviour that does not take into account the discovery of the new attribute. As a result of this, the □ parameter will be impossible to estimate. An estimable model requires situations where data exists on choices with and without the new attribute.

One alternative is to obtain a set of choice data from another location that includes the new attribute and transfer the values in a benefits transfer process. Alternatively, the *ex post* valuation of the attribute could be explored using revealed preference data after the discovery of the attribute. In neither case, however, can one tailor predictions of the effects of a new discovery to a specific area or site. The original revealed preference data must be augmented in this situation, since there is no market information about the effect of the new attribute on choice behaviour. One way to augment the revealed preference data is to add stated preference data. Stated preference data can be used to assess the change in intended behaviour that results from the introduction of new attribute. A potential solution to these problems is a combined analysis where revealed and stated preference information for the same set of individuals is pooled.

An empirical issue is the appropriate combination of the revealed and stated preference data. If only revealed data is used McFadden (1973) has shown that the choice probabilities take the form:

$$\exp \mu(V_j) / \sum_{k in C} \exp \mu(V_k),$$

where μ is a scale parameter. Since this parameter is not identified in a single set of data it is typically normalised to 1. Once the variables in the deterministic component of the indirect utility function, V, are specified and a functional form selected, the model becomes estimable using maximum likelihood methods.

When multiple data sets such as stated and revealed preference data are pooled an important issue is the consistency of the data sets with each other. A useful measure of this consistency is a test of the equality of the scale parameters in the two data sets. An econometric test of the equality of the scale parameters can be constructed. This is done by normalising *one* scale parameter in (2) and letting the scale parameters from the other data sets vary in the estimation process as shown by Swait and Louviere (1993), and Adamowicz et al. (1994; 1997). This method involves the notion that in any one data set μ is not identifiable, but that in any two (or more) datasets their ratio(s) can be identified (e.g. μ_1/μ_2). Thus, for a pooled revealed and stated preference data set this process involves the concatenation of the choice probabilities as follows:

$$\begin{aligned} RP: &\quad \pi(j) = \exp(\mu_{rp} V_j) / \sum_{k in C} \exp(\mu_{rp} V_k) \\ SP: &\quad \pi(j) = \exp(\mu_{srp} V_j) / \sum_{k in C} \exp(\mu_{sp} V_k), \end{aligned}$$

where μ_{rp} will be set to 1 and the μ_{sp} is a parameter to be estimated. Of course, this methodology can be used to extend the number of pooled data sets to any arbitrary size. If there were multiple new attributes one could extend the number of pooled data sets and concatenate choice probabilities.

In this analysis the method is applied to three data sets. The three data sets correspond to a single revealed and two stated preference data sets. The precise log likelihood function for this problem is given by:

$$LL = \sum_{n=1}^{N(RP)} \sum_{k in C} \ln \pi_n \{j \mid \beta\} + \sum_{n=1}^{N(SP_1)} \sum_{k in C} \ln \pi_n \{j \mid \beta, \mu_{sp1}\} + \sum_{n=1}^{N(SP_2)} \sum_{k in C} \ln \pi_n \{j \mid \beta, \mu_{sp2}\}.$$

The first part of the log likelihood corresponds to the revealed preference data while the second two pieces correspond to the two stated preference data sets. Note that there are three scale parameters, but that only two are estimated. These two are then compared to the normalised scale parameter to determine whether they are statistically different from one.

Data

The study involves wilderness recreation in Nopiming Provincial Park, Manitoba (Figure 1). The park is a 1440 km^2 area located about 145 km east of Winnipeg and is situated in the Precambrian or Canadian Shield. The area contains numerous rock outcrops that can rise as much as 36 m above the surrounding countryside and are a dominant feature. Pictographs are frequently found on rock outcrops along watercourses in this region, and while no paintings have been reported in the park, there are some in similar areas around Nopiming and in more remote areas in Ontario. The park has several river systems that contain small rapids and waterfalls and thus are attractive to backcountry recreationists interested in canoeing and kayaking. Most of the park is forested. Jack pine is the most abundant tree species in the park, although considerable areas of black spruce, aspen and white spruce can be found.

The wilderness recreation in this park and the surrounding region has been carefully studied in recent years. This involved an economic assessment of the importance of fire, forest ecosystems and other features (Boxall et al. 1996; Englin et al. 1996). As a result there is a detailed inventory of features along canoe routes that has been verified through intensive field work and GIS databases. The inventories identified areas in Nopiming that could potentially have rock paintings.

A registration system was developed to provide an understanding of the frequency of visitation to the backcountry areas of the park. In 1995 the registrants were surveyed. The survey included a stated preference experiment in which backcountry visitors were asked to respond to the possible presence or discovery of rock paintings in the park. The survey sample was created using the names of the leaders of the recreation parties who registered for a backcountry trip in Nopiming Park in 1993 or 1994. The original sample of 661 registrants was reduced to 587 by eliminating multiple trips by the same individual and incomplete addresses. The sample included individuals from 5 Canadian provinces, and 3 American states.

The experiment involved presenting pictures of two pictographs to respondents. The first involved a "pristine" pictograph. This pictograph exists in a more remote wilderness area in Ontario northeast of Nopiming. The second involved a picture of a pictograph located in a remote area in northern Manitoba that had been defaced by vandals and appears to be weathered. These pictures and the stated preference questions used in the experiment are shown in Figure 2.

The survey design exploits the knowledge of historical trip behaviour. Each respondent was offered the chance to change his or her trip to another route to see a rock painting. Since the original trip was known, each respondent was offered the rock paintings at a site they had not visited during the study period (1991-1994). Thus, the experiment ensured that every respondent had an opportunity to change his or her original site choice to a different site. The pictographs were offered at two routes: the Seagrim Lake canoe route and the Manigotagan River route. These sites were chosen because they had rock outcrops similar to those where pictographs are typically found.

The survey included a total of three mailouts. First, a questionnaire and cover letter was sent to the 587 individuals in early March 1995. Two weeks later a reminder post card was sent to any individual that had not responded. Finally, five weeks after the original mailout, a second questionnaire and cover letter was sent to nonrespondents. These procedures resulted in the return of 431 completed questionnaires which, adjusting for undeliverables (e.g. people moving etc.), represented a response rate of 81%.

The final data sets used for analysis consist of actual site choices for the respondents (revealed preference data) and their stated choices from the questionnaire (stated preference data). In this information, the choice set was limited to the eight major routes in the park. Any respondent whose actual trips were not to any of these eight routes was excluded from the analysis. This resulted in a final sample consisting of 386 respondents with complete trip data.

Results

The actual site choices of the respondents and their response to the SP experiment is shown in Table 1. Note that the Tulabi route was the most popular route actually chosen. About 42% of the respondents in the sample indicated they would change their actual route choice to another route to view a pristine painting. This change would occur regardless of the route where a painting was discovered (Seagrim and Manigotagan). However, only about 10% of the respondents would change their behaviour to view a defaced painting. The effect of the pictograph attributes on site choice is portrayed in Figure 3 where the cumulative increase in the number of trips to Seagrim and Manigotagan is shown relative to the availability of the two paintings.

The welfare measures associated with both the pristine and vandalised rock paintings are quantified in this analysis. In this analysis, the *a priori* hypothesis was that the pristine painting provides substantial positive benefits to the recreationists. This arises because the paintings enhance the attributes of some alternatives in the choice set. Thus, sites with paintings should exhibit an increased probability of visitation. It was further hypothesised that the vandalised paintings would not provide benefits as large as the pristine painting. However, defacement aside, the vandalised picture may still induce some change in trip behaviour by increasing the probability of visiting the sites with paintings.

Table 2 shows the parameters for six econometric models. The first three columns in the table report results for the individual revealed preference, pristine pictograph stated preference, and defaced pictograph stated preference models. Column 4 shows the results of the model combining revealed preference data with the pristine stated preference data and column 5 the results of combining revealed preference with defaced stated preference data. The last column shows the final model that includes the revealed preference data and **both** sets of the stated preference data.

In all of the models the parameters on the distance between an individual's home and the recreation site, hectares of recent burned areas, and hectares of black spruce old growth ecosystems are negative and significant. The parameters on hectares of white spruce growth and the single ASC for the Manigatogan canoe routes are positive and significant. These results are consistent with previous research on site choice behaviour in the park involving a larger sample of canoe routes (Boxall et al. 1996) and with trip data from different years (Englin et al 1996). In the RP model there is no parameter for pictographs because the paintings are not available. However, pictograph parameters are in the SP data. For the pristine pictograph model, the parameter on the picture is large and positive, while in the defaced pictograph model the parameter on the pictograph is smaller, but still positive. These findings are consistent with *a priori* expectations. In the three joint models the individual parameters are similar to those in the other models.

Tests of the equality of the restricted (joint) and unrestricted (single) models were conducted using likelihood ratio tests. These results are reported in Table 3. In each comparison the hypothesis of equality between models is not rejected at better than the 5% level of significance. In particular, the hypothesis of equality for the three-way joint model (RP+SPp+SPd) is not rejected. This means that the single RP and SP models share the same preference structures as the joint RP-SP models. Thus, unlike Adamowicz et al. (1994; 1997) it is not necessary to scale the SP data to the RP data. The ratios of scale parameters in these data are not significantly different than 1.0.

These specification tests support the use of the RP-SPp-SPd model to assess the welfare effects of discovering pictographs at the various routes in the park. Simulations were conducted to value the presence of the pristine and defaced pictographs at each of the eight major routes in the park. Figure 4 portrays the mean compensating variation per trip calculated over the sample using Hanemann's (1982) formula. At Seagrim and Manigotagan the presence of pristine pictographs would change site choice behaviour, providing benefits valued at $6.81 and $4.79 per trip respectively. These benefits would fall to $0.55 and $0.36 per trip if the painting was vandalised. Thus at these routes, a pristine pictograph would provide about 12-13 times the benefits of a vandalised one. At the other six routes the magnitudes of the benefits is higher or lower, but the pattern of the difference between the pristine and defaced paintings is similar. The overall magnitudes of the benefits across the sites reflect the complementarity of the pictographs with other attractive or negative features of the routes used in the choice models.

Discussion

A challenge facing managers of public lands is the tension between use, overuse and risk. A clear case in point are cultural resources such as the pictographs studied in this paper.

This analysis examined the value of pictographs to wilderness recreationists. In this study pristine pictographs are quite valuable, in some cases as much as $7.00 per trip. This compares favorably with museum admission charges. A defaced pictograph, however, is worth about a twelfth of the pristine pictograph. This contrast suggests that concern over the damages from vandalism is well founded. Of course, knowing the values of the pictographs does not solve the management conundrum. There remains the question of whether it is worthwhile forgoing the benefits associated with the pristine pictograph to reduce the risk of the pictograph being vandalized. Knowledge of the risk of vandalism is also needed to conduct a rational policy discussion. Nevertheless, without estimates of value however, no economic discussion of the merits of different policies can be conducted.

Early work joining stated and revealed preference data in random utility models struggled to develop methods that tested the consistency of the behavior suggested by revealed and stated data. Quite often the two data sets were not consistent with one another. In this analysis the two (actually three) data sets do support the hypothesis that the stated and revealed data come from consistent behavioral models. This finding is likely to have resulted from several factors. One is the clarity of the good in question. Pictographs are well known to Canadians who live in that region, and are certainly well known to those who visit wilderness areas there. Secondly, the population of canoeists is sufficiently homogeneous to make simple specifications of the scale parameter possible. A more heterogeneous population may not provide the scaling results seen in this study. Finally, this study was undertaken as part of a larger effort focused on modeling wilderness site choice behavior in the Canadian Shield region.

In this larger context the role of landscape features, ecosystem processes such as forest fires, and wilderness managerial features were understood. This knowledge helped to clarify the processes that are driving the choices of wilderness canoeists in the region. Furthermore, the detailed information base about the recreationists allowed the survey used in this study to be "custom designed" for each respondent. This design, in concert with the high level of knowledge of the factors affecting site choice behavior, may have contributed to the success of the modeling effort reported here.

References

Adamowicz, W.L., J. Louviere, and M. Williams. 1994. Combining stated and revealed preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management* 26:271-296.

Adamowicz, W.L., J. Swait, P.C. Boxall, J. Louviere, and M. Williams. 1997. Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. *Journal of Environmental Economics and Management* 32:65-84.

Boxall, P.C., D.O. Watson, and J. Englin. 1996c. Backcountry recreationists' valuation of forest and park management features in wilderness parks of the western Canadian Shield. *Canadian Journal of Forest Research* 26:982-990.

Cameron, T.A. 1992. Combining contingent valuation and travel cost data for the valuation of nonmarket goods. Land Economics 68(3):302-317.

Englin, J., P.C. Boxall, K. Chakraborty, and D.O. Watson. 1996. Valuing the impacts of forest fires on backcountry forest recreation. *Forest Science* 42(4):450-455.

Englin, J. and T. A. Cameron. 1996. Augmenting travel cost models with contingent behavior data. *Environmental and Resource Economics* 7(2):133-147.

Hanemann, W.M., 1982. Applied welfare analysis with qualitative response models, Working Pap. No. 241. Univ. California, Berkeley, 26 pp.

Maddala, F.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. New York: Cambridge University Press.

McFadden, D. 1973 Conditional logit analysis of qualitative choice behavior. In: P. Zarembka (Editor), *Frontiers in Econometrics*. Academic Press, New York, pp. 105-142.

Morey, E.R. 1994. "What is Consumer 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 Surplus?" Journal of Environmental Economics and Management. 26:257- 270.

Rajnovich, G. 1994.

Swait, J.R. and J.J. Louviere. 1993. The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research* 30:305-314.

Tables:

Table 1. Actual and hypothetical site choices in response to aboriginal pictographs by wilderness recreationists at Nopiming Provincial Park.

Routes	Original number of trips	Places where pictographs were offered	Number of people switching for a pristine pictograph	Number of people switching for a defaced pictograph
Tulabi	183			
Shoe	12			
Rabbit	58			
Seagrim	41	246	103	28
Gem	12			
Beresford	40			
Manigotagan 1	19	140	58	14
Manigotagan 2	21			
Total	386		161	42

Table 2. **Parameters (standard errors) from Conditional Logit Models used to Examine Recreation Site Choice in Nopiming Provincial Park, Manitoba.**

Variables	Single Models			Joint Models (Combined RP-SP)		
	RP	SPp	SPd			
	No Pic	Pristine Pic	Defaced Pic	RP + SPp	RP + SPd	RP+SPp+SPd
Distance	-0.0415*	-0.0364*	-0.0331*	-0.0420*	-0.0379*	-0.0384*
	(.0043)	(.0046)	(.0039)	(.0032)	(.0029)	(.0025)
Recent Burns	-0.2112*	-0.1499*	-0.1451*	-0.1921*	-0.1799*	-0.1743*
	(.0210)	(.0027)	(.0215)	(.0162)	(.0147)	(.0128)
Black Spruce	-1.3294*	-0.9455*	-0.8511*	-1.1685*	-1.0984*	-1.0481*
Old Growth	(.1394)	(.1555)	(.1321)	(.1038)	(.0961)	(.0814)
White Spruce	5.8763*	4.0251*	3.5704*	5.6880*	5.0427*	4.9443*
Old Growth	(.7410)	(.9015)	(.7304)	(.5897)	(.5325)	(.4620)
Good Pictograph		2.1462*		2.3629*		2.2703*
		(.1573)		(.1296)		(.1206)
Defaced Pictograph			0.3302*		0.4923*	0.4382*
			(.1892)		(.1772)	(.1735)
ASC-Man	2.8415*	2.7282*	2.1039*	3.1169*	2.5693*	2.7292*
	(.4843)	(.4974)	(.4286)	(.3505)	(.3272)	(.2726)
Log L	-666.31	-640.97	-717.40	-1307.38	-1383.45	-2027.02

Table 3. Hypothesis tests of parameter equality between the recreation site choice models.

Models	Log Likelihood	Likelihood Ratio Test	
		Δ Log L	χ^2
No Pictograph (RP)	-666.31		
Pristine Pictograph (SPp)	-640.97		
Defaced Pictograph (SPd)	-717.40		
RP + SPp	-1307.38	-1307.28	0.20
RP + SPd	-1383.45	-1383.71	0.52
RP + SPp + SPd	-2027.02	-2024.68	4.68 [1]

[1] critical χ^2 at P=0.05, 7 df is 14.07

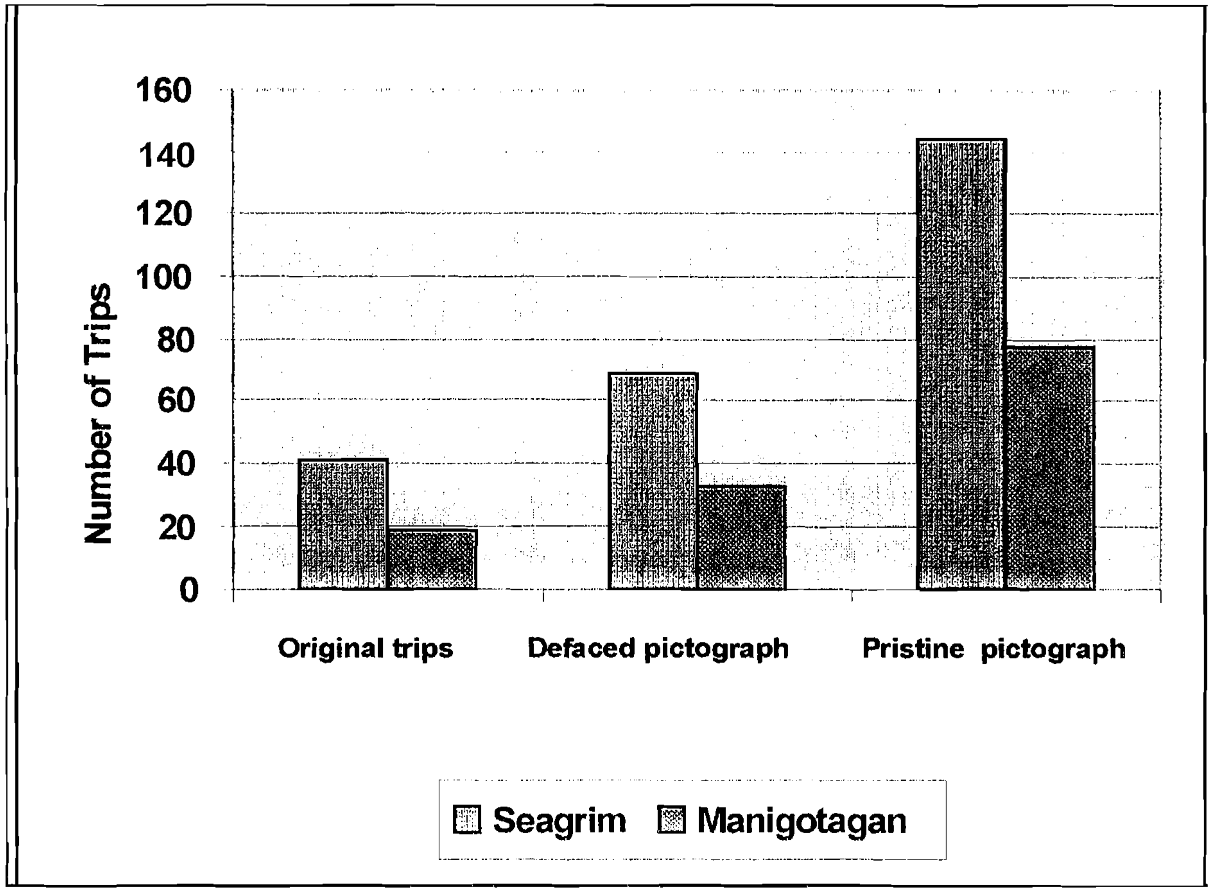


Figure 2. The effect of introducing pictographs of various types on the distribution of trips at two canoe routes in Nopiming Park.

A Comparison of Contingent and Actual Voting Behavior

Patricia A. Champ and Thomas C. Brown
Rocky Mountain Research Station
USDA Forest Service

1. Introduction

The 1992 National Oceanic and Atmospheric Administration (NOAA) panel, which was convened to assess the reliability of contingent valuation (CV) for estimating natural resource damages, recommended the CV question be posed as a vote on a referendum. Their recommendation was motivated by the incentive properties associated with the referendum mechanism, and the familiarity to respondents of the referendum format for making decisions about public goods. Along with this recommendation, external validation of the contingent values was highlighted as an important issue. The research reported in this paper compares contingent and actual voting behavior to investigate whether there is evidence of hypothetical bias associated with the contingent voting behavior.

II. Previous Studies

Previous comparisons of contingent and actual referenda have taken two forms. Polasky, Gainutdinova, and Kerkvliet (1996) and Carson, Hanemann, and Mitchell (1986) tried to predict actual voter referenda about public goods based on contingent referenda conducted just prior to the actual vote. Polasky, Gainutdinova, and Kerkvliet found that when 80 to 90% of the survey respondents who said they were undecided about how they would vote are coded as No votes, the hypothetical and actual referenda results are very similar. Likewise Carson, Hanemann and Mitchell found the survey results to be good predictors of the actual referendum outcome after adjusting for undecided voters.

Another approach to comparing actual and hypothetical referenda has been to conduct laboratory experiments. Cummings and Osborne (1996); Bjornstad, Cummings, and Osborne

(1997), and Cummings, Elliot, Harrison, and Murphy (1997) conducted laboratory experiments which involved implementing actual and hypothetical referenda to test for hypothetical bias. Four public goods are used across the various research projects. They find significant differences in the percent Yes respondents between the hypothetical and actual referendum for three of the four goods and conclude that these differences are due to hypothetical bias.

III. This Study

Like the first two studies mentioned above, this study compares responses to a contingent voting question with a subsequent actual referendum. Residents of Fort Collins, Colorado, were asked to vote on a referendum that was placed on the November 5, 1996, presidential election ballot. Voters were asked to vote for or against an ordinance that would allow the city to retain $764,000 in surplus revenue and use it for road maintenance. The wording to the referendum was as follows:

> An ordinance authorizing the city of Fort Collins (without increasing the current rate of city taxes) to retain, as a voter approved revenue change the sum of $764,000, and to spend that sum for the purposes of (1) constructing and repairing streets, (2) removing snow and de-icing streets, and (3) constructing sidewalks to promote pedestrian safety. This revenue was collected by the city in 1995 and is above the revenue and spending limits established under Article X, Section 20 of the Colorado Constitution.

During the week prior to November 5, a phone survey was conducted with a sample of Fort Collins residents randomly selected from the list of registered voters. In the survey, respondents were asked if they planned to vote in the November 5 election, in how many of the past four presidential elections they had voted, and if they had heard or read about the referendum concerned with whether the City of Fort Collins should retain surplus revenue collected in 1995. The referendum issue was read verbatim and respondents were asked how they would vote if the election were held today. The response to this question is what we compare to the actual election result. Respondents were also asked some attitude questions. Note that our study did not ask about a range of offer amounts as one would in a normal CV study. As a result, our analysis is restricted to comparing the proportion Yes in the real and contingent referenda.

The wording of the ordinance did not mention that if the referendum did not pass, the surplus funds would be refunded to Fort Collins residents on their utility bill.[1] If the referendum did not pass, each household would receive a refund of approximately $17. Although this information was included in newspaper articles about the referendum, few of our phone respondents were aware of the refund or the size of the refund.

After the election, a list of voters was purchased. From this list we were able to determine who in our sample had actually voted. We were also able to compare the characteristics of the phone survey respondents to all voters on the measures available in the voter list such as gender, year registered to vote in Larimer County, and party affiliation.

IV. Results

Despite the fact that calls were made only one week prior to the actual election, 41% of survey respondents said they hadn't heard of the referendum and 2% said they were not sure if they had heard of the referendum. The presidential election was of course the big news topic during this time but this referendum issue was covered in the local paper on more than one occasion. Almost all the respondents (98%) said they planned to vote or had already voted (via mail). Matching with the list of people who actually voted showed that only 63% of the survey

[1]The City of Fort Collins provides residents with electric service as well as water.

respondents voted on the November 5 ballot. However, a comparison of the survey respondents to the actual voting population, based on the measures (gender, year registered to vote in Larimer County, and party affiliation) available for the two groups, failed to find any significant differences between them.

The actual referendum passed with a 73% Yes vote. Seventy five percent of the survey respondents said they would vote Yes. Considering only those survey respondents that actually voted, 74% said they would vote Yes (Table 1). Actually voting did not seem to affect the distribution of responses to the contingent voting question (Table 2). The conservative approach to dealing with uncertain responses is to code them as No responses. Another option would be to exclude these respondents from the analysis. This approach is more controversial. As Table 2 shows, uncertain respondents are *not* less likely to vote, so exclusion seems inappropriate. Taking the conservative approach and re-coding the undecided votes as No, the responses of the survey respondents are similar to those of the actual voters. This result is consistent with the previous studies of this nature.

Contingency Table Analysis

To assess the construct validity of the contingent voting data, we examine relationships between other measures elicited in the phone survey and the response to the referendum question. Having heard of the referendum prior to the phone survey did not seem to affect the distribution of Yes and No responses.[2] However, respondents who hadn't heard of the referendum prior to the phone survey were more likely to respond "don't know" to the contingent voting question

[2]When the uncertain responses are removed from the data, the distribution of responses to the contingent voting question are not significantly different between those that had heard of the referendum prior to the phone survey and those that had not.

(Table 3). Survey participants were asked what they thought would happen to the $764,000 if the ordinance did not pass. Only 204 of the 531 (38%) phone respondents mentioned that the money would be refunded. The response to the contingent voting question is not independent of whether the refund was mentioned (Table 4). Twenty-four percent of the respondents who mentioned the refund said they would vote against the referendum. This compares to only 14% of those that did not mention the refund.

Fifty-four percent of the phone survey respondents were women. Women were more likely than men to vote in favor of the ordinance (Table 5).

The wording of the referendum mentions three uses of the surplus revenue (construct and repair streets, remove snow and de-ice streets, and construct sidewalks to promote pedestrian safety). Phone survey respondents were asked if they thought each of these was a good use of the surplus funds. Thinking a particular use of the funds is a good use is related statistically to the response to the contingent voting question (Tables 6, 7, 8). Despite the fact the many of the respondents were not aware of the referendum prior to the phone survey and the wording of the referendum was somewhat cryptic, the respondents seemed to be focusing on the details of the referendum when making the contingent voting decision.

The referendum wording did not mention the default situation if the referendum did not pass. Most of the survey respondents did not know that households would receive a refund if the referendum did not pass. During the phone interview, but after the initial contingent voting question, respondents were told that households would receive a refund if the referendum did not pass. Respondents were also asked about a question about each use of the surplus funds. After being told this information, respondents were asked again how they would vote on the referendum. Table 9 shows the effect of this information with uncertain responses to both questions are coded as No. Very few respondents who first voted Yes changed their vote to No. However, forty-six percent of the respondents who voted No to the initial question switched their vote to Yes once informed. It is interesting to note that the distribution of responses to the "informed" contingent voting question (86% Yes and 14% No) is less like the actual vote than the distribution of responses to the initial contingent voting question.

Multivariate Models

Logistic regression models (Table 10) suggest that removing the uncertain responses to the contingent voting question from the data analysis improves the fit of the model. However, as shown in Table 2, uncertain respondents are no less likely to actually vote, so removal of these respondents may not be appropriate.

V. Conclusions

Our goal was to investigate hypothetical bias associated with a contingent referendum. Our results are similar to the results of the studies which implemented similar research methods in that we do not find evidence of significant hypothetical bias. However, this research suggests a larger issue that needs to be addressed.

This study parallels a standard CV study in the sense that prior to receiving the phone call, many respondents were not very familiar with the good, and even if they were familiar, they were not well informed (i.e. they didn't know about the refund or the size of the refund). This study is very different from an actual CV study in the sense that a well developed CV survey would provide much more detail than does an actual referendum about the good or project of interest before asking respondents how they would vote.

We found that when we told phone survey respondents about the refund, the size of the refund, and asked about each of the proposed uses of the surplus revenue, and then asked them how they would vote on the referendum if it were today, the distribution of responses to that question was LESS like the actual vote than the "uninformed" distribution of responses to the contingent voting question. This result suggests that responses to a well developed CV question might differ from the "benchmark" actual referendum.

CV practitioners should think more about what it is they are trying to accomplish with a CV survey. If we are trying to create a situation similar to a real referendum, perhaps a well developed scenario with lots of information is not appropriate, as people are usually not so well informed when they actually vote in a referendum. However, if we want people to make fully informed decisions, perhaps actual referenda are not good benchmarks for assessing the external validity of contingent values.

REFERENCES

K.J. Arrow, R. Solow, P. Portney, E.E. Leamer, R. Radner, and H. Schuman. 1993. "Report of the NOAA Panel on Contingent Valuation" *Federal Register*, 58(10), January 15, 1993, 4602-4614.

D. Bjornstad, R. Cummings, and L. Osborne. Forthcoming. "A Learning Design for the Contingent Valuation Method" *Environmental and Resource Economics*.

R.T. Carson, W.M. Hanemann, and R.C. Mitchell. 1986. "The Use of Simulated Political Markets to Value Public Goods" Unpublished Manuscript. University of California, San Diego.

R.G. Cummings, S. Elliott, G.W. Harrison, and J. Murphy. Forthcoming. "Are Hypothetical Referenda Incentive Compatible?" *Journal of Political Economy*.

R.G. Cummings, and L.L. Osborne. 1996. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method" Working Paper, Policy Research Center, Georgia State University, Atlanta.

S. Polasky, O. Gainutdinova, J. Kerkvliet. 1996. "Comparing CV Responses with Voting Behavior: Open-space Survey and Referendum in Corvallis, Oregon" Proceedings from W-133 Benefits and Costs Transfer in Natural Resource Planning, Ninth Interim Report compiled by Joseph A. Herriges, Iowa State University.

Table 1:	**Actual Referendum and Survey Results**			
	N	**For**	**Against**	**Uncertain**
Actual Referendum	43,216	31,627 73%	11,589 27%	
Phone Survey Respondents	583	435 75%	94 16%	54 9%
Phone Survey Respondents who Voted	364	271 74%	57 16%	36 10%

Table 2:	**Relationship between whether actually voted and how vote if referendum was today (not statistically significant at 5% level)**		
	How vote if referendum was today?		
Voted in Nov. 96 election?	**For**	**Against**	**Uncertain/Don't Know**
Voted **Row %** **Column %**	271 74% 62%	57 16% 61%	36 10% 67%
Did not vote **Row %** **Column %**	164 75% 38%	37 17% 39%	18 8% 33%

Table 3:	**Relationship between whether heard of the referendum prior to the phone survey and how vote if referendum was today (statistically significant at 5% level)**		
	How vote if referendum was today?		
Heard of the referendum?	**For**	**Against**	**Uncertain/Don't Know**
Yes **Row %** **Column %**	243 75% 56%	61 19% 65%	20 6% 37%
No **Row %** **Column %**	183 74% 42%	33 13% 35%	30 12% 56%
Don't know **Row %** **Column %**	9 69% 2%	0	4 31% 7%

Table 4:	**Relationship between whether mention refund and how vote if referendum was today (statistically significant at 5% level)**	
	How vote if referendum was today?	
Mentioned Refund?	**For**	**Against (Uncertain coded as against)**
Did not mention **Row %** **Column %**	281 86% 65%	46 14% 48%
Mentioned **Row %** **Column %**	154 76% 35%	50 24% 52%

Table 5: Relationship between gender and how vote if referendum was today (statistically significant at 5% level)		
	How vote if referendum was today?	
Gender	For	Against (Uncertain coded as against)
Female	252	36
Row %	88%	12%
Column %	58%	38%
Male	182	60
Row %	75%	25%
Column %	42%	62%

Table 6: Relationship between thinking construction of sidewalks to promote pedestrian safety is good use of funds and how vote if referendum was today (statistically significant at 5% level)		
	How vote if referendum was today?	
Sidewalks good use of funds?	For	Against (Uncertain coded as against)
Yes	406	56
Row %	88%	12%
Column %	93%	58%
No	29	40
Row %	42%	58%
Column %	7%	42%

Table 7: Relationship between thinking removal of snow and de-icing of streets is a good use of surplus funds and how vote if referendum was today (statistically significant at 5% level)		
	How vote if referendum was today?	
Removal of snow and de-icing streets good use of funds?	**For**	**Against (Uncertain coded as against)**
Yes	388	59
Row %	87%	13%
Column %	89%	62%
No	47	37
Row %	56%	44%
Column %	11%	38%

Table 8: Relationship between thinking construction and repair of streets is a good use of surplus revenue and how vote if referendum was today (statistically significant at 5% level)		
	How vote if referendum was today?	
Construction and repair of streets is a good use of funds?	**For**	**Against (Uncertain coded as against)**
Yes	420	64
Row %	87%	13%
Column %	97%	67%
No	15	32
Row %	32%	68%
Column %	3%	33%

Table 9: Relationship between vote knowing about refund and initial vote (statistically significant at 5% level)			
	Response to initial vote question (no mention of refund)		
Response to vote question after told about refund	For	Against	Total
For Row % Column %	371 88% 99%	51 12% 46%	422 100% 86%
Against Row % Column %	5 8% 1%	61 92% 54%	66 100% 14%
Total[3] Row % Column %	376 77% 100%	112 23% 100%	488 100% 100%

[3]The distribution of responses is somewhat different than Table 2 because there was item non-response on the informed contingent vote question.

Table 10: Logistic Regressions		
	Model with uncertain responses coded as No (N=569)	Model with uncertain responses excluded (N=519)
Constant	5.9171 * (.8150)	7.7738* (1.0490)
Have you heard or read about the referendum? (1=yes; 2=no)	-.3914 (.2573)	-.0467 (.3270)
Mentioned that the city would refund money if ordinance not passed (1=didn't mention; 2=mentioned)	-.3914 (.2573)	-.7169* (.3178)
Gender (1=M; 0=F)	-.2100 (.2180)	-.8547* (.2821)
Do you think constructing sidewalks to promote pedestrian safety is a good use of funds? (1=yes; 2=no)	-1.2780* (.3083)	-1.6520* (.3425)
Do you think removing snow and de-icing streets is a good use of funds? (1=yes; 2=no)	-.2549 (.3118)	-.1800 (.3745)
Do you think constructing and repairing streets is a good use of funds? (1=yes; 2=no)	-1.8032* (.3746)	-2.2117* (.4380)
Log Likelihood	-277.75	-188.11
Percent predicted correctly	79.44%	86.13%

Dependent Variable = how would you vote if the referendum were today? 1=Yes; 0=No

Significantly different from zero at 5% level

Estimation and Welfare Calculations In a Generalized Corner Solution Model With An Application To Recreation Demand*

Daniel J. Phaneuf, Catherine L. Kling, and Joseph A. Herriges
Department of Economics, Iowa State University

JEL Classification: C25, Q26

* The authors would like to thank Arne Hallam, Yanis Bilias, Douglas Miller, and Noel Cressie for their helpful comments on earlier drafts of this paper. We are also indebted to Richard Bishop and Audrey Lyke, who gathered the Wisconsin Great Lakes angling data used in our analysis. As usual, any remaining errors are our own. This research was supported in part by the U.S. Environmental Protection Agency and by the Western Regional Research Project W-133. Although the research described in this article has been funded in part by the United States Environmental Protection Agency through R82-3362-010 to the authors, it has not been subject to the Agency's required peer review policy and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

Abstract

The Kuhn-Tucker model of Wales and Woodland (1983) provides utility theoretic framework for estimating preferences over commodities for which individuals often choose not to consume one or more of the goods. Due to the complexity of the model, however, there have been few applications in the literature and little attention has been paid to the problems of welfare analysis within the Kuhn-Tucker framework This paper provides an application of the model to the problem of recreation demand. In addition, we develop and apply a methodology for estimating compensating variation, relying on Monte Carlo integration to derive expected welfare changes.

I. INTRODUCTION

Traditional econometric methods for modeling consumer demand rely upon the specification of an indirect utility function, Roy's Identity, and the assumption of an interior solution to the consumer's utility maximization problem in order to derive an estimable system of demand equations. There are many applications, however, in which the assumption of an interior solution is unrealistic and, instead, corner solutions prevail. For example, in modeling recreation demand, it is typical to find that most households visit only a small subset of the available sites, setting their demand for the remaining sites to zero.[1] Similar corner solutions emerge in studies of both labor supply (e.g., Ransom (1987a, b), Lacroix and Fortin (1992), and Fortin and Lacroix (1994)) and food demand (e.g., Wales and Woodland (1983) and Yen and Roe (1989)).[2] In these situations, it is well known that failure to allow for the possibility of zero expenditure on one or more goods can lead to inconsistent estimates of consumer preferences.

Two broad strategies have emerged in the literature to deal with corner solutions. The first strategy, labeled the Amemiya-Tobin model by Wales and Woodland (1983), approaches the corner solutions dilemma from a statistical perspective. Systems of demand equations are initially derived without regard to non-negativity restrictions. The model then enforces these restrictions by employing an extension of Tobin's (1958) limited dependent variable model for single equations, later generalized by Amemiya (1974) for systems of equations. In particular, a truncated distribution for the random

---

[1] See Bockstael, Hanemann, and Strand (1986) and Morey, *et al.* (1995) for general discussions of non-participation and corner solution problems in the context of recreation demand.

disturbances is used to ensure non-negative expenditure shares, while allowing for a non-trivial proportion of the sample to have zero expenditure on one or more goods. Applications of the Amemiya-Tobin model have been implemented for a variety of goods. A sampling includes Wales and Woodland's (1983) analysis of meat demand and Heien and Wessells' (1990) study of general food consumption.

Such a statistical perspective has dominated the recreation demand literature. Single demand models or systems of demands for recreation have been estimated using a variety of estimators, including the tobit, Heckman, and Cragg models (Bockstael, Strand, McConnell, and Arsanjani (1990), Ozuna and Gomez (1994), Smith (1988), and Shaw (1988)), and a variety of count data models (Smith (1988), and Englin and Shonkwiler (1995)). Morey (1984) estimates a system of share equations that adopts a density function assuring strictly positive shares. The strand of this literature that has focused on multiple recreation sites has taken the Amemiya-Tobin model one step further. A two-stage budgeting argument has been used to separately analyze the total number of trips and the allocation of those trips among the available recreation sites.[3] The first stage site selection models use a discrete choice random utility framework. Corner solutions are then explicitly controlled for in the second stage model of the total number of trips using estimators that correct for censoring alone (Bockstael, Hanemann, and Kling (1989) and Morey, Shaw, and Rowe (1991)) or in combination with count models (Creel and Loomis (1990), Feather, Hellerstein, and Tomasi (1995), Hausman, Leonard,

---

[2]Corner solutions can also emerge for producers, both due to non-negativity constraints (e.g., Lee and Pitt (1987) and to upper bounds externally imposed by quotas (e.g., Fulginiti and Perrin (1993)).

and McFadden (1995), and Yen and Adamowicz (1994)). Although representing a range of estimation approaches, these models all share the Amemiya-Tobin reliance on statistical adjustments to represent corner solutions.

The second strategy for dealing with corner solutions takes a more structural or behavioral approach to the problem. Dubbed the Kuhn-Tucker model by Wales and Woodland (1983), it begins by assuming that individual preferences are randomly distributed over the population. The standard Kuhn-Tucker conditions characterizing individual utility maximization are then also randomly distributed, providing the basis for probabilistic statements regarding when corner conditions will occur and for constructing the likelihood function used in estimation. Initially developed by Wales and Woodland (1983) and Hanemann (1978) starting with the direct utility function, the approach has subsequently been extended to a dual form starting with the specification of the indirect utility function (Lee and Pitt (1986a) and Bockstael, Hanemann, and Strand (1986)). The appeal of the Kuhn-Tucker strategy lies in the unified and internally consistent framework it provides for characterizing the occurrence of corner solutions. However, due to the complexity of the model, there have been few applications (e.g., Wales and Woodland (1983), Lee and Pitt (1986b), Srinivasan and Winer (1994), and Ransom (1987a)) and none in the area of recreation demand.[4] Furthermore, little attention has been paid to the problem of welfare analysis within the Kuhn-Tucker framework. Due to

---

[3] See, for example, Bockstael, Hanemann, and Kling (1987), Hausman, Leonard, and McFadden (1995), Parsons and Kealy (1995), and Feather, Hellerstein, and Tomasi (1995).

[4] Morey, Waldman, Assane, and Shaw (1995) describe the Kuhn-Tucker model in the context of recreation demand, suggesting that it is the preferred approach, Bockstael, Hanemann, and Strand (1986) provide specifications appropriate for recreation demand, and Kling (1986) employs a form of the model to

the non-linearity of the model. closed form solutions for compensating or equivalent variation will typically not be available, requiring instead the use of Monte Carlo integration techniques.

The purpose of this paper is two-fold. First, we provide an empirical application of the Kuhn-Tucker model to the problem of recreation demand and site selection, modeling the demand for fishing in the Wisconsin Great Lakes region. Federal and state agencies are actively involved in management of the local fish populations and environmental conditions in this region. Understanding the demand for the resulting recreation opportunities will allow regulators to better evaluate existing programs and the impact of potential policy changes. Second, we develop and apply a methodology for estimating compensating variation in the context of the Kuhn-Tucker model, relying on Monte Carlo integration to derive expected welfare changes.

II. MODEL SPECIFICATION

A. Behavioral Model

The Kuhn-Tucker model begins with the assumption that consumers preferences over a set of M+1 commodities can be represented by a random utility function, which they maximize subject to a budget constraint and a set of non-negativity constraints. In particular, each consumer solves:

$$\underset{\mathbf{x},z}{Max}\; U(\mathbf{x},z,\mathbf{q},\gamma,\varepsilon) \tag{1}$$

s.t.

---

generate simulated data. However, none of these authors estimate the model or suggest how such a model could be used to compute welfare estimates.

$$\mathbf{p}'\mathbf{x} + z \leq y \tag{2a}$$

and

$$z \geq 0, x_j \geq 0, j = 1, \ldots, M \tag{2b}$$

where $U(\cdot)$ is assumed to be a quasi-concave, increasing, and continuously differentiable function of $(\mathbf{x}, z)$. $\mathbf{x} = (x_1, \ldots, x_M)'$ is a vector of goods to be analyzed, z is the numeraire good, $\mathbf{p} = (p_1, \ldots, p_M)'$ is a vector of commodity prices, y denotes income, and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_M)'$ is a vector of random disturbances capturing the variation in preferences in the population. The disturbance vector is assumed to be known to the individual, but unobservable by the analyst. The vector $\mathbf{q} = (q_1, \ldots, q_M)'$ represents attributes of the M commodities.[5] The inclusion of commodity attributes is particularly important in recreation demand studies since policy analysis is often interested in the welfare implications of changing the environmental quality of a site.

The first-order necessary and sufficient Kuhn-Tucker conditions for the utility maximization problem are then given by:

$$U_j(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) \equiv \frac{\partial U(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon)}{\partial x_j} \leq \lambda p_j, x_j \geq 0, x_j U_j(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) = 0 \quad j = 1, \ldots, M, \tag{3a}$$

$$U_z(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) \equiv \frac{\partial U(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon)}{\partial z} \leq \lambda, \; z \geq 0, \; zU_z(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) = 0, \tag{3b}$$

and

$$\mathbf{p}'\mathbf{x} + z \leq y, \lambda \geq 0, (y - \mathbf{p}'\mathbf{x} - z)\lambda = 0, \tag{3c}$$

[5] In general, a vector of attributes may characterize each commodity. However, we have used a scalar attribute here to simplify notation.

where λ denotes the marginal utility of income. For simplicity, we assume that the numeraire good is a necessary good, so that equation (3b) can be replaced by

$$\lambda = U_z(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon). \tag{3b'}$$

In addition, since $U(\cdot)$ is increasing in $\mathbf{x}$ and z, the budget constraint will be binding, with

$$z = y - \mathbf{p}'\mathbf{x}. \tag{3c'}$$

Substituting equations (3b′) and (3c′) into (3a) yields the M first-order conditions associated with the commodities of interest:

$$U_j(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, \varepsilon) \le p_j U_z(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, \varepsilon), \; x_j \ge 0, x_j\left[U_j - U_z p_j\right] = 0 \;\; j = 1, \dots, M. \tag{3a'}$$

Finally, we assume that $U_{z\varepsilon} = 0$, $\partial U_j / \partial \varepsilon_k = 0 \; \forall k \ne j$, and $\partial U_j / \partial \varepsilon_j > 0 \; \forall j = 1, \dots, M$, so that[6]

$$U_j(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, \varepsilon) = U_z(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, \varepsilon) p_j, \; j = 1, \dots, M \tag{4}$$

defines a set of implicit equations for the ε_j's of the form:

$$\varepsilon_j = g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma), \; j = 1, \dots, M \tag{5}$$

and the first-order conditions in equation (3a′) can be rewritten as:

$$\varepsilon_j \le g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma), \; x_j \ge 0, \; x_j g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma) = 0 \quad j = 1, \dots, M. \tag{6}$$

Equation (6), along with the specification of the joint density function $f_\varepsilon(\varepsilon)$ for ε, provides the necessary information to construct the likelihood function for estimation. Consider an individual who chooses to consume positive quantities for only the first k

---

6 Wales and Woodland (1983) accomplish this by assuming that the errors enter the utility function such that $U_j(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) = \overline{U}_j(\mathbf{x}, z; \mathbf{q}, \gamma) + \varepsilon_j \; j = 1, \dots, M$. See Bockstael, Hanemann, and Strand (1986) and Morey *et al.* (1995) for more general treatments of the error term.

commodities (i.e., $x_j > 0$, $j = 1,\ldots,k$ and $x_j = 0$, $j = k+1,\ldots,M$). Their contribution to the likelihood function is given by the probability

$$\int_{-\infty}^{g_{k+1}} \cdots \int_{-\infty}^{g_M} f_\varepsilon(g_1,\ldots,g_k,\varepsilon_{k+1},\ldots,\varepsilon_M)\, abs\left|J_k\right|\, d\varepsilon_{k+1}\cdots d\varepsilon_M \tag{7}$$

where J_k denotes the Jacobian for the transformation from ε to $(x_1,\ldots,x_k,\varepsilon_{k+1},\ldots,\varepsilon_M)'$. There are 2^M possible patterns of binding non-negativity constraints for which a probability statement such as (7) can be constructed. The likelihood function can then be formed as the product of the appropriate probabilities and maximum likelihood used to recover estimates of the utility function's parameters.

B. Conditional Utility Functions and the Computation of Welfare Effects

A primary reason for estimating the structure of consumer preferences over a set of commodities is to provide a basis for welfare analysis. In particular, policymakers may be interested in the welfare implication of changing the price or quality characteristics of the existing set of alternatives, or of reducing the number of alternatives available. Formally, let $V(\mathbf{p}, y; \mathbf{q}, \gamma, \varepsilon)$ denote the solution to the utility maximization defined in equations (1) and (2) above. The compensating variation (C) associated with a change in the price and attribute vectors from $(\mathbf{p}^0, \mathbf{q}^0)$ to $(\mathbf{p}^1, \mathbf{q}^1)$ is implicitly defined by

$$V(\mathbf{p}^0, y; \mathbf{q}^0, \gamma, \varepsilon) = V(\mathbf{p}^1, y + C(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma, \varepsilon); \mathbf{q}^1, \gamma, \varepsilon). \tag{8}$$

There are several important attributes of the compensating variation measure that are worthy of note. First, from the analyst's perspective, $C(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma, \varepsilon)$ is a random variable. Policy makers will typically be interested in the average value of this measure in the population, $\overline{C}(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma)$. Second, the non-linearity of the utility maximization problem will typically preclude a closed form solution for C or its average. As a result, numerical techniques will be required.[7]

The process of computing C can be clarified by considering the utility maximization as a two-stage process, in which the individual maximizes his or her utility conditional on a set of binding non-negativity constraints and then chooses among the resulting conditional indirect utility functions.[8] Formally, let

$$A = \{\varnothing, \{1\}, \ldots, \{M\}, \{1,2\}, \{1,3\}, \ldots, \{1,2, \ldots, M\}\} \tag{9}$$

denote the collection of all possible subsets of the index set $I = \{1, \ldots, M\}$. A conditional indirect utility function $V_\omega(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon)$ can then be defined for each $\omega \in A$ as the maximum utility level the consumer can achieve when they are restricted to the commodities indexed by ω. Formally:

$$V_\omega(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon) = \underset{\mathbf{x},z}{Max}\, U(\mathbf{x}, z, \mathbf{q}, \gamma, \varepsilon) \tag{10}$$

s.t.

$$\sum_{j \in \omega} p_j x_j + z \leq y \tag{11a}$$

and

$$z \geq 0, x_j = 0, j \notin \omega, \; x_j \geq 0 \; j \in \omega, \tag{11b}$$

where $\mathbf{p}_\omega = \{p_j : j \in \omega\}$ is the vector of commodity prices that have not been constrained to zero. Let $\mathbf{x}_\omega(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon)$ denote the conditional demand levels solving this utility

---

[7] This problem is similar to the one encountered in nonlinear site selection models and recently addressed by McFadden (1995) and Herriges and Kling (1996).

maximization problem. Notice that, since the prices associated with those commodities that have been forced to zero do not enter the budget constraint in (11a), V_ω and $\mathbf{x}_\omega$ are both functions of $\mathbf{p}_\omega$ and not $\mathbf{p}$. However, both the conditional indirect utility function and conditional demand equations will depend on the entire vector of quality attributes, $\mathbf{q}$, and not simply $\mathbf{q}_\omega = \{q_j : j \in \omega\}$, unless the property of weak complementarity is imposed (Maler, 1974).[9]

Constraining a subset of the commodities to have zero consumption provides, of course, no assurance that the optimal consumption levels for the remaining commodities will be positive. Let

$$\widetilde{A} \equiv \widetilde{A}(\mathbf{p}, y; \mathbf{q}, \gamma, \varepsilon) = \{\omega \in A : x_{\omega j}(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon) > 0, \forall j \in \omega\} \tag{12}$$

denote the collection of ω's for which the corresponding conditional utility maximization problem yields an interior solution. The original consumer utility maximization problem can then be viewed as a two-stage problem in which conditional indirect utility functions are computed for each $\omega \in A$ and then the consumer chooses the V_ω that maximizes his or her utility. That is[10]

---

[8] Hanemann (1984) originally detailed this argument in the case of extreme corner solutions (i.e., when only one of the commodities is consumed). Bockstael, Hanemann, and Strand (1986) extend the argument for the general case.

[9] Imposing weak complementarity implies that there is only "use value" associated with the commodities. In the absence of weak complementarity, individuals may also assign "non-use" value to a commodity (i.e., the individual perceives utility from the availability of a good without actually consuming it). Here, we adopt Freeman's (1993) definitions of use, non-use, and existence values and note, as an aside, that models based on observed behavior cannot elicit information on existence value.

[10] The second equality follows from the fact that for all $\omega \notin \widetilde{A}$, the associated conditional utility maximization problem yields a binding non-negativity constraint for some $j \in \omega$. The solution is, therefore, redundant, being equivalent to another utility maximization problem (defined by $\widetilde{\omega} \subset \omega$ with $\widetilde{\omega} \in \widetilde{A}$) where that good has been constrained to zero *a priori*.

$$V(\mathbf{p},y;\mathbf{q},\gamma,\varepsilon) = \underset{\omega \in A}{Max}\{V_\omega(\mathbf{p}_\omega,y;\mathbf{q},\gamma,\varepsilon)\} = \underset{\omega \in \tilde{A}}{Max}\{V_\omega(\mathbf{p}_\omega,y;\mathbf{q},\gamma,\varepsilon)\}. \tag{13}$$

The computation of the compensating variation in equation (8) then corresponds to implicitly solving for $C(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma,\varepsilon)$ in

$$\underset{\omega \in \tilde{A}^0}{Max}\{V_\omega(\mathbf{p}^0_\omega,y;\mathbf{q}^0,\gamma,\varepsilon)\} = \underset{\omega \in \tilde{A}^1}{Max}\{V_\omega(\mathbf{p}^1_\omega,y+C(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma,\varepsilon);\mathbf{q}^1,\gamma,\varepsilon)\}. \tag{14}$$

Notice that the index collection $\tilde{A}$ may change as a result of the changing price and/or quality attribute levels.[11]

There are three difficulties associated with computing $\overline{C}(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma)$ in practice. First, for any given ε and γ, $C(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma,\varepsilon)$ is an implicit function for which no closed form solution typically exists. However, numerical procedures, such as numerical bisection, can be readily applied to solve this problem. Second, given $C(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma,\varepsilon)$ and γ, $\overline{C}(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma)$ does not have a closed form solution. However, Monte Carlo integration can be used, resampling from the underlying distribution of ε, $f_\varepsilon(\varepsilon)$, and averaging $C(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma,\varepsilon)$ over the draws of ε.[12] Third, given an algorithm for computing $\overline{C}(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma)$, the analyst does not typically have available γ, but instead must rely upon an estimator $\hat{\gamma} \sim g_{\hat{\gamma}}$ (e.g., the maximum likelihood estimator of γ). Thus, any computation of $\overline{C}$ will itself be a random variable, dependent upon the distribution of $\hat{\gamma}$. We employ the procedure developed by Krinsky and Robb (1986) to approximate the statistical properties of $\hat{C}$, our estimate of $\overline{C}$, repeatedly drawing

[11] Policy changes may also involve the elimination of initially available sites. Such changes can be reflected in the make-up of the index collection $\tilde{A}$.

[12] See Geweke (1996) for a useful review of Monte Carlo integration.

realizations from $g_{\hat{\gamma}}$ and computing $\hat{C}$ for each of these realizations. Formally, the above elements are combined into the following numerical algorithm:

- A total of N_γ parameter vectors (i.e., $\gamma^{(i)}$, $i = 1, \dots, N_\gamma$) are randomly drawn from the distribution $g_{\hat{\gamma}}$.

- For each $\gamma^{(i)}$ and each observation in the sample ($n = 1, \dots, N$), a total of N_ε vectors of random disturbance terms (i.e., $\varepsilon^{(ink)}$, $k = 1, \dots, N_\varepsilon$) are randomly drawn from the distribution $f_\varepsilon(\varepsilon)$.

- Substituting $\gamma^{(i)}$ and $\varepsilon^{(ink)}$ for γ and ε in equation (14), numerical bisection can then be used to solve for C, with the result labeled $C^{(ink)}$.

- Averaging $C^{(ink)}$ over the N_ε draws from the disturbance distribution and the N observations in the sample yields $\hat{C}^{(i)}$, a Monte Carlo integration evaluation of $E_\varepsilon\left[C(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma^{(i)}, \varepsilon)\right]$.

- The distribution of $\hat{C}^{(i)}$'s provides the basis for characterizing the distribution of the mean compensating variation of interest ($\overline{C}$) in light of our uncertainty regarding γ. The mean value of $\hat{C}^{(i)}$ over the N_γ parameter draws provides a consistent estimate of $\overline{C}$. The distribution of the $\hat{C}^{(i)}$'s can be used to construct standard errors for our estimate of $\overline{C}$.

C. Empirical Specification

In our application below, we employ the empirical specification suggested by Bockstael, Hanemann, and Strand (1986). In particular, we assume that the consumer's direct utility function is a variant of the linear expenditure system, with

$$U(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) = \sum_{j=1}^{M} \Psi_j(q_j, \varepsilon_j) \ln(x_j + \theta) + \ln(z) \tag{15}$$

and

$$\Psi_j(q_j, \varepsilon_j) = \exp\left(\sum_{s=1}^{S} \delta_s q_{js} + \varepsilon_j\right) \quad j = 1, \dots, M \tag{16}$$

where $\gamma = (\delta, \theta)$ and q_{js} denotes the s^{th} quality attribute associated with commodity j. The Ψ_j's can be thought of as quality indices associated with each good. The parameter θ provides an indication of whether there is non-use value associated with the commodities being modeled, since weak complementarity holds in the above model only if we restrict $\theta = 1$.

One advantage of the above utility function is that the implicit equations for the ε_j's in equation (4) that result from the Kuhn-Tucker conditions can be explicitly solved, yielding the following equivalent first-order conditions:

$$\varepsilon_j \leq g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma), \quad x_j \geq 0, \quad x_j g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma) = 0 \quad j = 1, \ldots, M. \tag{17}$$

where

$$g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma) = \ln\left[\frac{p_j(x_j + \theta)}{y - \sum_{j=1}^{M} p_j x_j}\right] - \sum_{s=1}^{S} \delta_s q_{js} \quad j = 1, \ldots, M. \tag{18}$$

Specifying a joint distribution for the random disturbances (i.e., $f_\varepsilon(\varepsilon)$) completes the empirical model. We assume that the ε_j's are independent and identically distributed negative extreme value variates with parameters $\eta = 0$ and λ. An important feature of this specification is that closed form equations exist for the probabilities in the likelihood function. In particular, the probability of observing the usage pattern $\omega \in \tilde{A}$ is given by:

$$\exp\left(-\sum_{j\in\omega} \frac{g_j}{\lambda}\right) abs|J_\omega| \exp\left[-\sum_{j=1}^{M} \exp\left(\frac{-g_j}{\lambda}\right)\right] \tag{19}$$

where J_{ω} denotes the Jacobian of the transformation from ε to

$\left\{\varepsilon_j \text{ for } j \notin \omega \text{ and } x_j \text{ for } j \in \omega\right\}$.[13]

III. DATA

Our empirical application of the Kuhn-Tucker model focuses on angling in the Wisconsin Great Lakes region. The data are drawn primarily from two mail surveys of angling behavior conducted in 1990 by Richard Bishop and Audrey Lyke at the University of Wisconsin-Madison.[14] The surveys provide detailed information on the 1989 angling behavior of Wisconsin fishing license holders, including the number and destination of fishing trips to the Wisconsin Great Lakes region, the distances to each destination, the type of angling preferred, and the socio-demographic characteristics of the survey respondents. A total of 509 completed surveys were available for analysis, including 266 individual who had fished the Wisconsin Great Lakes region for lake trout or salmon and 247 who fished only inland waters of Wisconsin (i.e., non-users from the perspective of the Great Lakes region). While the surveys provide data on 22 distinct Great Lake fishing destinations, we have combined these destinations into four aggregate "sites":

- Site 1: Lake Superior
- Site 2: South Lake Michigan
- Site 3: North Lake Michigan, and
- Site 4: Green Bay.

---

13 The Jacobian transformations J_{ω}, while not difficult, are algebraically messy and relegated to

This aggregation divides the Wisconsin portion of the Great Lakes into distinct geographical zones consistent with the Wisconsin Department of Natural Resources' classification of the lake region.

The price of a single trip to each of the four fishing sites consists of two components: the cost of getting to the site (i.e.. direct travel cost) and the opportunity cost of the travel time. Round trip direct travel costs were computed for each destination and each individual by multiplying the number of round trip miles for a given individual-destination combination by the cost per mile for the vehicle class driven, as provided by the American Automobile Association. The cost of the travel time was constructed using one-third of the individual's wage rate as a measure of the hourly opportunity cost of recreation time and assuming an average travel speed of forty-five miles per hour to compute travel time.[15] The price of visiting a destination p_j is then the sum of the direct travel cost and the cost of the travel time.

Two types of quality attributes (i.e., q_{js}'s) are used to characterize the recreation sites: fishing catch rates and toxin levels. Catch rates are clearly important site characteristics since the anticipated success of fishing is likely to be a major determinant in the recreation decision. Furthermore, state and federal agencies currently spend large amounts of time and money to influence catch rates in the region through stocking programs and regulations. The inclusion of catch rates as a quality attribute in the model

---

Appendix A.

[14] Details of the sampling procedures and survey design are provided in Lyke (1993).

[15] There is an extensive debate on appropriate measure of the opportunity cost of travel time. Since it is not a purpose of this study to enter into this debate, we have chosen this relatively simple means of accounting for the travel time cost, drawing on research results of McConnell and Strand (1981).

will allow it to be used to conduct welfare analyses of existing and/or alternative fishery management programs.

In constructing the catch rate variables, we focus our attention on the catch rates for the four aggressively managed salmonoid species: lake trout, rainbow (or steelhead) trout, Coho salmon, and Chinook salmon. Creel surveys by the Wisconsin Department of Natural Resources provide 1989 catch rates for each of these species at each of the 22 disaggregate destinations used in the angling surveys. Furthermore, these catch rates are broken down by angling method, including private boat, charter fishing, and pier/shore angling. Data from the Wisconsin angling survey were used to match the mode-specific catch rates to each individual fisher based upon their most frequent mode of fishing.

We include toxin levels as an additional quality attribute of each site since the presence of environmental contaminants is likely to influence the recreation decision and they provide a proxy for the overall level of water quality at the site (De Vault *et al.* (1996)). Toxins are found in varying levels in fish, water, and sediments throughout the Great Lakes and are routinely responsible for health warnings in the regions. De Vault *et al.*(1989) provide a study of toxin levels in lake trout during the relevant time period, with samples taken from locations throughout the Great Lakes. We use the average toxin levels (ng/kg-fish) from this study, matched on the basis of proximity to our four aggregate sites, to form a basic toxin measure T_j $(j = 1,\ldots,4)$ for each site.[16] However, toxin levels are likely to influence visitation decisions only if the consumer perceives that

[16] While there are a variety of toxins reported in the De Vault *et al.*(1989) study, we use the levels of toxins 2,3,7, 8-TCDD, which are generally responsible for the fish consumption advisories issued by states in the region.

the toxins create a safety issue. The Wisconsin angling survey asked respondents if the toxin levels in fish were of concern to them. We use this information to form an "effective toxin level" variable $E_j = T_j D$ $(j = 1,\ldots,4)$, where D = 1 indicates that the respondent was concerned about the toxin levels in fish and D = 0 otherwise.

With both catch rates and toxins included as quality variables, the quality index terms from equation (16) become

$$\Psi_j(\mathbf{q}_j, \varepsilon_j) = \exp\left[\delta_0 + \delta_{lk} R_{lk,j} + \delta_{ch} R_{ch,j} + \delta_{co} R_{co,j} + \delta_{rb} R_{rb,j} + \delta_E E_j + \varepsilon_j\right],\ j = 1,\ldots,4, \quad (20)$$

where $R_{k,j}$ denotes the catch rate for species k and site j, with $k=lk$ for lake trout, ch for Chinook salmon, co for Coho salmon, and rb for rainbow trout.

Tables 1 and 2 provide summary statistics for the data. Table 1 focuses on the mean and standard deviation of the usage, price, and quality characteristics for the four sites used in our analysis. Table 2 characterizes the trip usage patterns (i.e., ω) found in the Wisconsin angling survey data. Note that, while many (72%) of the visitors to the Great Lakes sites visit only one of the sites, a substantial percentage (28%) visit more than one site. Thus, neither an extreme corner solution (Hanemann (1984)) nor an interior solution model could accurately depict this group of consumers' choices.

IV. RESULTS

A. Model Estimation

The Kuhn-Tucker model of Wisconsin Great Lakes angling was estimated using maximum likelihood, yielding the parameter estimates provided in Table 3. All of the parameters have the expected signs and, with the exception of the coefficient on lake trout catch rates, are statistically different from zero at a 5% critical level or less. For example, one would expect, and we find, that higher toxins reduce the perceived quality of a site (i.e., $\delta_E < 0$). On the other hand, higher catch rates should enhance site quality (i.e., $\delta_k > 0$). This is the case for each of the fish species considered. Furthermore, the small and statistically insignificant coefficient on lake trout is not unexpected, since among anglers lake trout are typically considered a less desirable species. The other salmon species have a "trophy" status not shared by lake trout. In addition, the eating quality of lake trout is generally considered inferior to that of other species.

The other coefficient of direct interest in Table 3 is θ. Recall from Section II that the parameter θ provides a means of testing for the presence of non-use value. The assumption of weak complementarity, which is used extensively in the recreation demand literature and precludes non-use value, holds in our model only if $\theta = 1$. The results in Table 3 indicate that this restriction is not borne out in the current application. In particular, θ is statistically different from "1" using a 1% critical level, suggesting that some non-use value is associated with the four Great Lakes angling sites. We pursue this further in the welfare analysis below.

B. Welfare Analysis

One of the motivations for estimating models of recreation demand is to provide policy makers with estimates of the welfare implications of changing environmental quality or site availability. A primary advantage of the Kuhn-Tucker models is that it permits the construction of these welfare estimates in an internally consistent and utility-theoretic framework. The model simultaneously predicts changes to the sites visited and the total number of trips taken, which in turn determines changes in consumer utility. In this subsection, we use the estimated Kuhn-Tucker model in Table 3, along with the numerical procedures developed above, to evaluate a series of policy scenarios for the Wisconsin Great Lakes region.

The Great Lakes region provides many opportunities for policy-relevant welfare experiments as the lakes are heavily managed. The fishery itself is, in many ways, artificially created and maintained. Of the major species included in the model, only lake trout are native to both Lake Superior and Lake Michigan. Rainbow trout were introduced around the turn of the century, while the salmon species were not present until the 1950's. These species now reproduce naturally in the lakes, but are heavily augmented with stocking programs. The lakes have also been invaded by exotic species, including the sea lamprey. A parasite accidentally introduced in the 1930's, the sea lamprey decimated lake trout populations in the lakes. Efforts to reintroduce naturally reproducing lake trout to Lake Superior have been successful, while in Lake Michigan the population is completely maintained through stocking. Expensive sea lamprey control efforts continue to this day. Finally, there are ongoing efforts throughout the Great Lakes region to improve the fisheries by reducing the level of toxins entering the food chain from commercial and industrial sources. For each of these forms of intervention, the natural policy question arises as to whether the benefits of these programs are sufficient to offset the corresponding costs. Our Kuhn-Tucker model can be used to assess program benefits. As an illustration of this capability, we estimate welfare loss under three policy scenarios:

- Scenario A: Loss of Lake Michigan Lake Trout. Under this first policy scenario, state and local efforts to artificially stock lake trout in Lake Michigan and Green Bay would be eliminated. It is assumed that this would drive lake trout catch rates

($R_{lk,j}$) to zero for sites 2, 3, and 4, since the species is only naturally reproducing in Lake Superior (site 1).[17]

- Scenario B: Loss of Lake Michigan Coho Salmon: Under this policy scenario, state and local efforts to artificially stock Coho salmon in Lake Michigan and Green Bay would be suspended. Again, it is assumed that the corresponding Coho catch rates ($R_{co,j}$) would be driven to zero for sites 2, 3, and 4.[18]

- Scenario C: Reduced Toxin Levels. Under the final policy scenario, we consider the welfare implications of a twenty percent reduction in toxin levels (i.e., $E_j,\ j = 1,2,3,4$).

For each of these scenarios, mean compensating variation ($\overline{C}$) was estimated using GAUSS and the procedures outlined in Section IIB above. In particular,

- A total of $N_\gamma = 250$ parameter vectors (i.e., $\gamma^{(i)},\ i = 1,\ldots,N_\gamma$) were randomly drawn from the asymptotically justified normal distribution for the maximum likelihood parameter estimates $\hat{\gamma}$ in Table 3.

- For each $\gamma^{(i)}$ and each observation in the sample ($n = 1,\ldots,509$), a total of $N_\varepsilon = 1000$ vectors of random disturbance terms (i.e., $\varepsilon^{(ink)},\ k = 1,\ldots,N_\varepsilon$) were formed by drawing four independent extreme value variates.[19]

- Substituting $\gamma^{(i)}$ and $\varepsilon^{(ink)}$ for γ and ε in equation (14), numerical bisection was then used to solve for C, with the result labeled $C^{(ink)}$.

- Averaging $C^{(ink)}$ over the N_ε draws from the disturbance distribution and the N observations in the sample yields an estimate ($\hat{C}^{(i)}$) of the mean compensating variation for the j[th] draw from the estimated parameter distribution.

---

17 Under this scenario, it is assumed that the catch rate for lake trout in Lake Superior is unchanged, either because of ongoing stocking programs or the natural replenishment capabilities of the fishery.

18 Coho salmon do, in fact, naturally reproduce in Lake Michigan and Green Bay, so that the elimination of stocking programs would not drive the associated catch rates completely to zero. However, we use $R_{co,j} = 0$ to approximate the impact of dropping the stocking programs and to make comparisons to Scenario A more direct.

19 Simulations were used to determine that $N_\varepsilon = 1000$ was sufficient to reduce the standard deviation of $\hat{C}^{(i)}$ to less than two percent of $\hat{C}^{(i)}$.

The distribution of the $\hat{C}^{(i)}$'s provides the basis for characterizing the distribution of the mean compensating variation of interest ($\overline{C}$) in light of our uncertainty regarding the parameter estimates in Table 3. The mean value of the $\hat{C}^{(i)}$ over the 250 parameter draws provides a consistent estimate of $\overline{C}$ and are reported in column two of Table 4 for each scenario, with the corresponding standard deviations reported in parentheses.[20]

The total compensating variations in Table 4 have the expected signs and relative magnitudes, given the parameter estimates in Table 3. As expected, the loss of Coho salmon (Scenario B) has a greater impact on consumer welfare than the loss of lake trout (Scenario A). In particular, an average of almost $275 per angler per season would be required to compensate for the loss of Coho salmon in the Lake Michigan and Green Bay sites, whereas less than $40 per angler per season would compensate for the loss of lake trout in the same region. Furthermore, the lake trout benefits are not statistically different from zero using any reasonable confidence level, whereas the Coho benefits are significant at a 5% critical level. The lake trout results are particularly interesting from a policy perspective, since so much effort has gone into rehabilitating the lake trout fishery during the past three decades.

Turning to Scenario C, we find that a twenty percent reduction in toxin levels would have a substantial and statistically significant impact on angler welfare. Anglers would be willing to pay, on average, almost $75 per season for such a reduction.

---

20 Some caution should be exercised in using the standard deviations to construct confidence intervals. The $\hat{C}^{(i)}$'s are unlikely to be symmetrically distributed and, hence, two-standard deviation confidence intervals will be inappropriate. While the construction of asymmetric confidence intervals is conceptually straightforward, a substantially larger N_γ would be needed to precisely construct the necessary tail statistics (See, e.g., Efron and Tibshirani (1993)).

An interesting feature of our model in the context of welfare measurement is that it does not impose the property of weak complementarity as most models of recreation demand do. Rather, the model allows us to test for this property through the estimated value of θ, thus effectively testing for the existence of non-use value. As our estimate rejects weak complementarity, with $\hat{\theta}$ significantly greater than one, the total welfare measures reported in Table 4 for our three scenarios are comprised of both use and non-use components. An interesting question, then, is what portion of the compensating variations in Table 4 are due purely to use values?

To answer this question, we first isolate the non-use component by setting the prices of the sites high enough so that use is choked off at the relevant sites.[21] We then follow the procedures outlined above for $\overline{C}$. The resulting welfare measures are thus entirely associated with non-use of the resource. By subtracting these non-use values from the total values reported in column two of Table 4, we obtain estimates of the use value. These use values are provided in column 3 of Table 4. In all of these cases, the use value comprises roughly two-thirds of the total value.

V. SUMMARY AND CONCLUSIONS

In this study, we have provided an empirical application of the Kuhn-Tucker model to the problem of recreation demand, estimating the demand for fishing in the Wisconsin Great Lakes region as well as welfare measures associated with changes in site catch rates and toxin levels. The Kuhn-Tucker model is appealing for use in recreation

[21] See Appendix B for a formal discussion of the division of the total compensating variation between use and non-use values.

demand modeling in that it deals with the abundant observation of general corner solutions in an internally consistent and utility theoretic framework. The same model drives both the site selection choice and the total number of trips taken by recreationists. This feature is particularly important to the task of assessing welfare changes.

In our application to the Great Lakes region, we estimate the lost value to anglers of eliminating lake trout from Lake Michigan and Green Bay, the loss of Coho Salmon from Lake Michigan and Green Bay, and the welfare improvements associated with reduced toxin levels in the lakes. We present point estimates of both the pure use value of these changes and non-use value. In addition to providing point estimates of these welfare measures, we provide information on the reliability of the estimates in the form of standard errors.

An additional novel feature of the Kuhn-Tucker model estimated here is that we empirically test for the existence of the often imposed property of weak complementarity. Our model estimates reject the property, implying that there is non-use value associated with the resource. To our knowledge this model provides the first empirical test for the existence of non-use value using behaviorally based models (many applications of survey methods such as contingent valuation have presented results suggesting the existence of non-use values).

There are two areas where improvements to the model estimated here could be made. First, it would be desirable to explore alternative functional forms in the specification of individual utility. The trade-off here, of course, is in identifying forms that are both flexible and yet yield Kuhn-Tucker conditions that generate closed-form probabilities for the likelihood function. Second, it would be desirable to experiment with error distributions other than the extreme value to investigate the robustness of the results to the assumed error structure.

APPENDIX A: JACOBIAN TRANSFORMATIONS

The Jacobian transformations terms in equation (19) are given by:

$$J_\omega = F_j \text{ for } \omega = \{j\},\ j = 1,2,3,4. \tag{A1}$$

$$J_\omega = \prod_{j\in\omega} F_j - \prod_{j\in\omega} z_j \text{ for } \omega = \{1,2\},\{1,3\},\{1,4\},\{2,3\},\{2,4\}, \text{ and } \{3,4\}. \tag{A2}$$

$$J_\omega = \prod_{j\in\omega} F_j + 2\prod_{j\in\omega} z_j - \sum_{j\in\omega} F_j\left(\prod_{\substack{k\in\omega\\ k\neq j}} z_k\right) \text{ for } \omega = \{1,2,3\},\{1,2,4\},\{1,3,4\}, \text{ and } \{2,3,4\}. \tag{A3}$$

and

$$\begin{aligned} J_\omega = {} & F_1F_2F_3F_4 - 3z_1z_2z_3z_4 + 2(F_1z_2z_3z_4 + F_2z_1z_3z_4 + F_3z_1z_2z_4 + F_4z_1z_2z_3) \\ & -(F_1F_2z_3z_4 + F_1F_3z_2z_4 + F_1F_4z_2z_3 + F_2F_3z_1z_4 + F_2F_4z_1z_3 + F_3F_4z_1z_2) \text{ for } \omega = \{1,2,3,4\} \end{aligned} \tag{A4}$$

where

$$z_j \equiv \frac{p_j}{y - \sum_{k=1}^{4} p_k x_k} = \frac{\partial \varepsilon_k}{\partial x_j}\ \forall\ k \neq j, \tag{A5}$$

$$F_j \equiv \frac{1}{x_j + \theta} + z_j = \frac{\partial \varepsilon_j}{\partial x_j}\ \forall\ j, \tag{A6}$$

APPENDIX B: COMPUTING THE USE COMPONENT OF COMPENSATING VARIATION

The purpose of this appendix is to formally describe the reasoning behind the decomposition of total compensating variation into "use" and "non-use" components as presented in Table 4. In order to simplify the exposition, we abstract from the general corner solution problem by assuming an interior solution. The generalization to cases in which corner solutions emerge is straightforward, but tedious, and adds nothing to the intuition. In addition, we simplify the presentation by considering a simple two good problem.

Consider the problem of measuring the total compensating variation associated with changing the attribute of site 2 from q_2^0 to q_2^1, without changing the corresponding characteristics of site 1. This total compensating variation can be expressed in terms of expenditure functions as

$$C_{total} = e(p^0, q_1^0, q_2^0, U^0) - e(p^0, q_1^0, q_2^1, U^0), \tag{B1}$$

where p_j^0 denotes the initial cost of visiting site j and U^0 denote the individual initial level of utility. In the absence of weak complementarity, this total value can be divided into two components: use value and non-use value. We adopt the following definition of non-use value.

$$C_{non-use} = e(p_1^0, p_2^\bullet, q_1^0, q_2^0, U^0) - e(p_1^0, p_2^\bullet, q_1^0, q_2^1, U^0). \tag{B2}$$

where $p_j^\bullet$ denotes the choke price associated with site j. Thus, we define non-use value to be the compensating variation an individual consumer places on the change in environmental quality when the consumer does not consume any of the good whose quality changes. To derive an expression for the resulting use value it is only necessary to subtract non-use value in (B2) from the total value in (B1) yielding

$$\begin{aligned} C_{use} &= C_{total} - C_{non-use} \\ &= e(p_1^0, p_2^0, q_1^0, q_2^0, U^0) - e(p_1^0, p_2^0, q_1^0, q_2^1, U^0) - \\ &\quad e(p_1^0, p_2^\bullet, q_1^0, q_2^0, U^0) - e(p_1^0, p_2^\bullet, q_1^0, q_2^1, U^0) \\ &= \int_{p_2^0}^{p_2^\bullet} h_2(p_1^0, p_2, q_1^0, q_2^1, U^0)\,dp_2 - \int_{p_2^0}^{p_2^\bullet} h_2(p_1^0, p_2, q_1^0, q_2^0, U^0)\,dp_2. \end{aligned} \tag{B3}$$

As is clear from the expression in the third line of (B3), this definition of use value corresponds to the sum of the areas under the Hicksian demand curves for the good whose quality changes.

VI. REFERENCES

Amemiya, T., "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables are Truncated Normal." *Econometrica* 42 (September 1974), 999-1012.

Bockstael, N.E., W.M. Hanemann, and C.L. Kling, "Estimating the Value of Water Quality Improvements in a Recreation Demand Framework," *Water Resources Research* 23 (May 1987), 951-60.

Bockstael, N.E., W.M. Hanemann, and I.E. Strand, *Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models*, Vol. 2, Washington D.C.: U.S. Environmental Protection Agency, Office of Policy Analysis (1986).

Bockstael, N.E., I.E. Strand, K.E. McConnell, and F. Arsanjani, "Sample Selection Bias in the Estimation of Recreation Demand Functions: An Application to Sportfishing," *Land Economics* 66 (1990), 40-49.

Creel, M., and J.B. Loomis, "Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California, *American Journal of Agricultural Economics* 72 (May 1990), 434-45.

De Vault, D.S., R. Hesselberg, P.W. Rodgers and T.J. Feist. "Contaminant Trends in Lake Trout and Walleye from the Laurentian Great Lakes," *Journal of Great Lakes Research* 22 (1996), 884-895.

De Vault, D.S., D. Dunn, P. Bergqvist. K. Wiberg and C. Rappe, "Polychlorinated Dibenzofurans and Ploychlorinated Dibenzo-p-Dioxins in Great Lakes Fish: A Baseline and Interlake Comparison," *Environmental Toxicology and Chemistry* 8 (1989), 1013-1022.

Efron, B., and R.J. Tibshirani, *An Introduction to the Bootstrap*, New York: Chapman and Hall, 1993.

Englin, J., and S. Shonkwiler, "Estimating Social Welfare Using Count Data Models: An Application to Long-Run Recreation Demand Under Conditions of Endogenous Stratification and Truncation," *Review of Economics and Statistics* 77 (1995), 104-112.

Feather, P., D. Hellerstein, and T. Tomasi, "A Discrete-Count Model of Recreation Demand," *Journal of Environmental Economics and Management* 30 (1996), 316-322.

Fortin, B. and G. Lacroix, "Labor Supply, Tax Evasion and the Marginal Cost of Public Funds - An Empirical Investigation," *Journal of Public Economics* 55 (1994), 407-431.

Freeman, A. M., *The Measurement of Environmental and Resource Values: Theory and Methods*, Washington, D.C.: Resources for the Future, 1993.

Fulginiti, L., and R. Perrin, "The Theory and Measurement of Producer Response Under Quotas," *Review of Economics and Statistics* 75 (1993), 97-106.

Geweke, J., "Monte Carlo Simulation and Numerical Integration," in H. Amman, D. Kendrick, and J. Rust, eds., *Handbook of Computational Economics* (Amsterdam: Elsevier Science B.V., 1996).

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., "Discrete-Continuous Models of Consumer Demand," *Econometrica* 52 (1984), 541-561.

Hausman, J.A., 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," *Journal of Public Economics* 56 (Winter 1995), 1-30.

Heien, D., and C. R. Wessells, "Demand Systems Estimation with Microdata: A Censored Regression Approach," *Journal of Business & Economic Statistics* 8 (1990), 365-371.

Herriges, J. A., and C. L. Kling, "Nonlinear Income Effects in Random Utility Models," working paper, Iowa State University, October 1996.

Kling, C. L. *Measuring the Recreational Benefits of Environmental Amenities Using Multiple Site Models: An Evaluation of Techniques*, Ph.D. dissertation, University of Maryland, 1986.

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

Lacroix, G., and B. Fortin, "Utility-Based Estimation of Labor Supply Functions in the Regular and Irregular Sectors," *Economic Journal* 102 (1992), 1407-1422.

Lee, L.F., and M.M. Pitt, "Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach." *Econometrica* 54 (November 1986a), 1237-42.

Lee, L. F., and M.M. Pitt, "Specification and Estimation of Consumer Demand with Many Binding Non-Negativity Constraints," Center for Economic Research Discussion Paper No. 236. Department of Economics, University of Minnesota, October, 1986.

Lee, L.F., and M.M. Pitt. "Microeconometric Models of Rationing, Imperfect Markets, and Non-Negativity Constraints." *Journal of Econometrics* 36 (1987), 89-110.

Maler, K.. *Environmental Economics*. Baltimore: Johns Hopkins. 1974.

McConnell, K.E.. and I.E. Strand. "Measuring the Cost of Time in Recreation Demand Analysis: An Application to Sport Fishing," *American Journal of Agricultural Economics* 63 (1981), 153-156.

McFadden, D., "Computing Willingness-to-Pay in Random Utility Models." working paper, Department of Economics. University of California. Berkeley, December 1995.

Morey, E.R., "The Choice of Ski Areas: Estimation of a Generalized CES Preference Ordering with Characteristics, Quadratic Expenditure Functions. and Non-Additivity," *Review of Economics and Statistics* 66 (1984), 221-36.

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), 181-201.

Morey, E.R., D. Waldman. D. Assane. and D. Shaw, "Searching for a Model of Multiple-Site Recreation Demand that Admits Interior and Boundary Solutions." *American Journal of Agricultural Economics* 77 (1995), 129-140.

Ozuna. T., and I. Gomez, "Estimating a System of Recreation Demand Functions Using a Seemingly Unrelated Poisson Regression Approach," *Review of Economics and Statistics* 76 (1994), 356-60.

Parsons, G.R., and M.J. Kealy, "A Demand Theory for Number of Trips in a Random Utility Model of Recreation." *Journal of Environmental Economics and Management* 29 (1996), 418-433.

Ransom, M.R., "The Labor Supply of Married Men - A Switching Regressions Model," *Journal of Labor Economics* 5 (1987a), 63-75.

Ransom. M.R., "An Empirical Model of Discrete and Continuous Choice in Family Labor Supply," *Review of Economics and Statistics* 69 (1987b), 465-472.

Shaw, D. "On-Site Sample' Regression: Problems of Non-negative Integers. Truncation, and Endogenous Stratification." *Journal of Econometrics* 37 (1988), 211-23.

Smith, V.K., "Selection and Recreation Demand." *American Journal of Agricultural Economics* 70 (1988), 29-36.

Tobin, J., "Estimation of Relationships for Limited Dependent Variables. *Econometrica* 26 (1958), 24-36.

Wales, T.J., and A.D. Woodland. "Estimation of Consumer Demand Systems with Binding Nonnegativity Constraints." *Journal of Econometrics* 21 (1983), 263-285.

Yen, S. T., and W. L. Adamowicz. "Participation, Trip Frequency, and Site Choice: A Multinomial Poisson Hurdle Model of Recreation Demand." *Canadian Journal of Agricultural Economics* 42 (1994), 65-76.

Yen, S.T., and T.L. Roe. "Estimation of a 2-Level Demand System with Limited Dependent-Variables," *American Journal of Agricultural Economics* 71 (1989), 85-98

Table 1. – Average Site Characteristics (Standard Deviations in Parentheses)

	Lake Superior	North Lake Michigan	South Lake Michigan	Green Bay
1989 Fishing Trips (x_j)	2.75 (13.33)	1.56 (6.32)	2.35 (8.92)	0.65 (3.07)
Price (p_j)	177.84 (172.59)	123.70 (172.92)	85.88 (139.62)	129.11 (173.54)
Lake Trout Catch Rate ($R_{lk,j}$)	.046 (.059)	.022 (.030)	.029 (.045)	.001 (.002)
Chinook Salmon Catch Rate ($R_{ch,j}$)	.010 (.014)	.048 (.030)	.027 (.024)	.036 (.032)
Coho Salmon Catch Rate ($R_{co,j}$)	.028 (.021)	.005 (.005)	.040 (.053)	.005 (.008)
Rainbow Trout Catch Rate ($R_{rb,j}$)	.001 (.001)	.018 (.026)	.012 (.013)	.001 (.002)
Effective Toxin Level (E_j)	.597 (.491)	2.270 (1.866)	3.464 (2.847)	2.270 (1.866)

Notes: Catch rates are measured in terms of fish per person-hour of effort.

Table 2. – Distribution of Trips

Sites Visited	Number of Observations
All four sites. $\omega=\{1,2,3,4\}$	3
Lake Superior. North and South Lake Michigan. $\omega=\{1,2,3\}$	1
Lake Superior. North Lake Michigan. and Green Bay. $\omega=\{1,2,4\}$	7
Lake Superior. South Lake Michigan. and Green Bay. $\omega=\{1,3,4\}$	0
North and South Lake Michigan and Green Bay. $\omega=\{2,3,4\}$	13
Lake Superior and North Lake Michigan. $\omega=\{1,2\}$	10
Lake Superior and South Lake Michigan. $\omega=\{1,3\}$	8
Lake Superior and Green Bay, $\omega=\{1,4\}$	2
North and South Lake Michigan. $\omega=\{2,3\}$	13
North Lake Michigan and Green Bay. $\omega=\{2,4\}$	19
South Lake Michigan and Green Bay, $\omega=\{3,4\}$	4
Lake Superior. $\omega=\{1\}$	49
North Lake Michigan. $\omega=\{2\}$	46
South Lake Michigan. $\omega=\{3\}$	85
Green Bay, $\omega=\{4\}$	11
No sites visited. $\omega=\emptyset$	243

Table 3. – Parameter Estimates

Parameter	Estimate	P-Value
δ_0 (Intercept)	-8.53	$<.001$
δ_{lk} (Lake Trout)	0.10	.953
δ_{ch} (Chinook Salmon)	13.39	$<.001$
δ_{co} (Coho Salmon)	3.12	.023
δ_{rb} (Rainbow Trout)	8.61	.035
δ_E (Effective Toxin Level)	-0.06	.018
θ	1.76	$<.001$
λ	1.29	$<.001$
$\theta-1$	0.76	$<.001$

Table 4. – Welfare Estimates (Standard Errors in Parentheses)

Policy Scenario	Mean Compensating Variation ($\overline{C}$) Total	Use Only
Scenario A: Loss of Lake Trout Species at Sites 2, 3 and 4	39.78 (143.05)	28.28 (97.56)
Scenario B: Loss of Coho Salmon at Sites 2, 3 and 4	274.18* (123.18)	186.51* (80.79)
Scenario C: A 20% Reduction in Toxins at all Sites	-74.76** (26.13)	-50.31** (17.07)

Notes: A single asterisk indicates significance at the 5% level, while two asterisks indicate significance at the 1% level.

Estimating Willingness to Pay For Protecting Critical Habitat For Threatened And Endangered Fish With Respondent Uncertainty

Earl R. Ekstrand
Economics Group
Technical Service Center
U.S. Bureau of Reclamation
Denver, CO 80225

and

John Loomis
Department of Agricultural and Resource Economics
Colorado State University
Fort Collins, CO 80523

Acknowledgments: This research has benefited from discussions with Robert Berrens, University of New Mexico who generously shared his approach to this same problem. We appreciate the willingness of Patricia Champ, USDA Forest Service to share her preliminary analysis of uncertainty with us. Kelly Giraud of Colorado State University was instrumental in the survey focus groups, pretesting, and survey mailing.

Disclaimer: The opinions e2pressed here do not necessarily reflect the policy or views of the U.S. Bureau of Reclamation.92 422

This work was funded in part by the U.S. Bureau of Reclamation's General Investigation Program, project number 100-6520-0008-086-01-0-1-9.

Abstract

A comparison of the standard contingent valuation model to alternative modifications that explicitly incorporate respondent uncertainty is performed to estimate economic benefits of protecting critical habitat for nine threatened and endangered fish species living in the Colorado, Green and Rio Grande rivers. The standard dichotomous choice contingent valuation model estimated a value of $195 per household which was compared to values ranging from $28 to $247 depending on how respondent uncertainty was explicitly incorporated into the dichotomous choice model. For this dataset, incorporating respondent uncertainty had the effect of reducing the goodness of fit and increasing the variance of estimated willingness to pay.

Introduction

The Endangered Species Act (ESA) requires federal agencies including the Bureau of Reclamation (BOR) to avoid jeopardizing the continued existence of threatened and endangered (T&E) species, including fish species in the four corners region. Recovery measures comprise re-operation of dams, installation of fisheries protection measures and in-stream flow releases. These actions result in direct costs and opportunity costs of reduced irrigation and hydropower benefits. These economic losses are easy to see and quantify and tend to be concentrated among a small number of water users who publicize their losses.

Economic benefits, however, are more difficult to measure but people have shown they value the preservation of a wide variety of threatened and endangered species from the obscure striped shiner (a fish in the Milwaukee River) to the bald eagle and whooping crane, as summarized by Loomis and White (1996). While values per household may be quite low for some species, the public good nature of preserving endangered species result in large aggregate values as millions of households throughout the U.S. can simultaneously enjoy the benefits of knowing these species still exist. However, the dispersed nature of the public good benefits provide much less of an incentive for beneficiaries of preservation of endangered species to become actively engaged in the policy process.

Economic Benefits vs. Impacts

Most of the economic analyses related to species listing and critical habitat decisions have focused upon the short run effects on local jobs and incomes. This type of analysis is frequently called economic impact or regional economic effects analysis and often has little to do with the long run benefits or costs of species preservation. While these figures sometimes have significant shock value and influence in the affected region, rarely is it acknowledged that decreases in commodity production in one region are usually offset by increases in production (and corresponding employment gains) in other regions or other industries.

Society often realizes real opportunity costs from protecting T&E species and their habitats in the form of higher costs of production or valuable uses foregone. As such, economic benefits must be defined and measured in a commensurate fashion by managing agencies. Measuring benefits using willingness to pay (WTP) is the conceptually correct measure of benefits (Just, et al. 1982) and is the currently accepted norm among Federal agencies for benefit-cost analysis (U.S. Water Resources Council 1983) and natural resource damage assessment (Department of Interior 1986). Since the public owns T&E species, willingness to accept for avoiding losses would often be the more appropriate measure for estimating benefits. However the public's unfamiliarity with being offered compensation as compared to being asked to pay for programs, coupled with difficulties in empirical measurement, results in nearly all studies using WTP as the measurement technique. The reliance on a conservative measure such as WTP may help to off-set the concern that the survey technique used to elicit WTP (discussed below) may overstate values due to the hypothetical nature of the payment.

Contingent Valuation Method

Existence of threatened and endangered species is not a product that is sold in markets, but has value to society. Because of lack of price, economists have developed a hypothetical market method, called the Contingent Valuation Method (CVM), that uses a survey to measure household WTP to protect a species in a particular location. A CVM survey is a standardized and widely used method for obtaining WTP and involves developing a hypothetical market or referendum as a vehicle by which an individual reveals his or her WTP. While there are legitimate concerns about the degree of accuracy of CVM estimates of WTP for natural resources the public is unfamiliar with, CVM has been shown in empirical test-retest studies to be reliable (Kealy et al. 1988, Loomis 1989, 1990, and Carson et al. 1997).

CVM is recommended for use by Federal agencies for performing benefit-cost analysis (U.S. Water Resources Council 1983), for valuing natural resource damages (U.S. Department of Interior 1986), and was upheld by the Federal courts (U.S. District Court of Appeals 1989). Recently, a "blue ribbon panel" including two Nobel laureate economists, an environmental economist and a survey research specialist concluded that CVM can produce estimates reliable enough to be the starting point for administrative and judicial determinations (Arrow et al. 1993).

Previous Research on the Economic Value of T&E Species

Loomis and White (1996) provide a review of estimates of economic benefits for about 20 T&E species, about half not published. Their meta analysis of the values suggests that most of the variability in values of species can be explained by a few specific variables. In particular, the value per household is largely determined by the size of the change in species population being offered in the survey, whether visitors or households are being surveyed, whether the species is a bird, and whether annual or a one-time willingness to pay amount was being asked. Using both a linear and a double log functional form, the regressions explained between 58% and 68% of the variation in per household WTP. This high explanatory power for a cross-sectional study is encouraging regarding the internal consistency of CVM-derived WTP values. One important problem, though, ignored in valuation of T&E species to date is omission of the issue of respondent uncertainty. We now turn to that issue.

Past Research on Respondent Uncertainty

Most individuals are not familiar with many T&E species and have no prior experience paying for species protection. Many individuals realize personal satisfaction from knowing these species exist, but have not devoted much time contemplating how much they would pay. If they spent the time to reflect on the tradeoffs between household costs and preservation of species, they could refine their preferences. However, the one-shot nature of CVM survey responses may not provide sufficient repetition for generating stable preferences.

While CVM may not provide the opportunity to stabilize preferences in this way, respondents can express the level of confidence in their dollar bids and this information can be incorporated into the statistical analysis. Those individuals who have extensive prior knowledge of the environment or species in question may have well defined preferences and great certainty in their responses while those with little or no knowledge may have less defined preferences and therefore more uncertainty about their answers. Incorporating the stated uncertainty of respondents into the statistical model could improve the estimation and accuracy of the analysis (Manski 1995).

Several approaches have been recently developed to incorporate respondent uncertainty into CVM. Ready et al. (1995) used a polychotomous choice question format where the respondent had a choice of six responses to a single bid amount: definitely yes, probably yes, maybe yes, maybe no, probably no, and definitely no. They found that allowing for uncertainty increased WTP. Welsh and Bishop (1993) multiple bounded approach provided a similar range of responses for the full range of bid amounts. They found little change in the level of WTP, but the estimates had reduced variability.

A different approach was employed by Li and Mattsson (1996), using a two-step approach. They first used a conventional dichotomous choice WTP question followed by having the respondent perform a "post-decisional" rating of the certainty of the response to the WTP question. This certainty rating is incorporated into the likelihood function directly. The net result reduced both the mean WTP and the variance of the estimated WTP.

Champ et al. (1997), Johannesson et al. (1997), and Polasky et al. (1996) show the importance of addressing respondent uncertainty on WTP estimates. Champ et al. compared actual contributions of a sample of Wisconsin residents to remove roads on the North Rim of the Grand Canyon with stated WTP from a separate sample of Wisconsin residents. The stated WTP of $79 was several times the actual payment of $9. Respondents to the stated WTP questionnaire were also asked how certain they were of their response using a ten point scale where 10 was very certain and 1 was very uncertain. Champ et al. recoded all YES responses with certainty response of less than a 10 to a NO. This recoding technique led to reduction in the estimated WTP to about $12, very similar to the actual mean payment amount of $9. This provides strong evidence in favor of incorporating uncertainty into the analysis.

Johannesson et al. estimated WTP for a box of chocolates using a similar technique and found that recoding resulted in statistical under-estimation of the cash WTP. If only completely certain YES responses were retained as YES answers, the authors concluded that too many responses were recoded NO. The difference between the conclusions of Champ et al. and Johannesson et al. may be due to potential for free riding in the public good in contrast to the private good study.

Polasky et al. performed a different type of validity study, comparing intended voting behavior in an actual referendum for open space. While there was not a perfect match of characteristics between voters and survey respondents, the results shed light on the issue of dealing with uncertain voters. The actual referendum had about 44.8 percent voting YES. Excluding those individuals who were uncertain and those refusing to answer, the various sample frames answering the CVM question all yielded 53- 54 percent YES responses, significantly different from the actual vote. When the uncertain respondents and those refusing were coded NO, the percentage of YES votes dropped to about 40 to 43 percent, slightly understating the actual vote pattern. Polling literature support this result suggesting that most undecided voters choose NO in the actual vote (Magelby 1989).

The purpose of our paper is to develop new methods for more fully utilizing information on response uncertainty and compare these to recently proposed approaches for incorporating uncertainty into statistical models for estimating WTP. We compare the existing and new methods in terms of variance of mean WTP and goodness of fit of the logit model.

Measuring and Incorporating Respondent Uncertainty

Our research more fully utilizes the information contained in the 1-10 post-decisional rating that respondents provide regarding the certainty of the willingness to pay response. The first three models extend the Champ et al. and Polasky et al. approach of recoding uncertain YES responses as NOs by evaluating three different cut-off points for a YES response to count as a yes and created three models called Yes10, Yes910, and Yes810. Based on our scale of 1 being very uncertain and 10 being very certain, Yes10 recoded all YES responses to NOs if the respondent did not have a certainty response of 10, following the suggestion of Champ et al. To this we add a Yes910 model which recoded all YES responses that were not a certainty of 9 or 10, and the YES810 recoded those not having a certainty of 8, 9, or 10 to NOs.

The next three models were designed to explicitly incorporate more of the uncertainty information into the logistic regression model. Our fourth model continued to follow the suggestion of recoding only YES responses scaling these according to respondent uncertainty and can be visualized on a certainty scale as follows:

|-----------|-----------|-----------|-----------|-----------|-----------|-----------|----------|-----------|-------
----|
0 .1 .2 .3 .4 .5 .6 .7 .8 .9 1.0
NO <---Increasing uncertainty of YES

This model is appropriate if respondents answering NO are certain they would not pay while the YES responses are more uncertain. Since YES responses are coded as "1" and NOs coded as "0", multiplying this variable by the certainty level (converted to probability) replaces the YES responses with a scale of .1 to 1 while all NO responses remain a zero. This model can be directly estimated in LIMDEP (Greene 1992) and we refer to it as the asymmetric uncertainty model, or ASUM. The logit model is:

(1) $$\text{Prob (YES)} = 1-\{1+\exp [B_0 - B_1 (\$X)]\}^{-1},$$

where \$X is the dollar amount the individual is asked to pay and B_0 and B_1 are the intercept and slope coefficients respectively.

In our last two models, we assume that the YES and NO responses are equally uncertain and therefore re-scale both the YES and the NO answers. We used two approaches for re-scaling. The first approach for recoding both YES and NO WTP responses is similar to the approach by Li and Mattsson, overlaying the YES and NO certainty levels on the same scale. This requires re-scaling both the YES and NO responses between 0 and 1.0, which we refer to as the LIMATT model. Their approach assumes that "... a yes response with 40% confidence, for example, is equivalent to a no answer with 60% confidence" (Li and Mattsson, 1995:264). This is shown as:

|----------|-----------|------------|-----------|-----------|-----------|-----------|----------|----------|---------
--|
0 .1 .2 .3 .4 .5 .6 .7 .8 .9 1.0
<---Increasing uncertainty of
YESs
Increasing uncertainty of NOs --
>

Our second approach gives greater weight to the actual response to the WTP question. That is, the YES answers are re-scaled only between .5 and 1.0 depending on certainty level while the NO respondents are re-scaled only between 0 and .5. The probabilities then are arrayed along a continuum as follows:

|-----------|----------|-----------|-----------|-----------|-----------|-----------|----------|-----------|--------
---|
0 .1 .2 .3 .4 .5 .6 .7 .8 .9 1.0
-->< --
increasing uncertainty of NO response increasing uncertainty of YES response

This Symmetric Uncertainty Model (referred to as SUM) can also be estimated with LIMDEP using logistic regression. It retains the individual level data using the recoded probabilities directly. The SUM suggests a structure which uses the uncertainty to modify the strength of the YES or NO response but retains the respondent's YES or NO answer. We believe this is more consistent with literal interpretation of the survey answers than Li and Mattsson's approach.

Comparisons of Approaches

Li and Mattsson as well as Manski suggest that with more information incorporated into the statistical model, goodness of fit will increase and the WTP estimates will be more precise. We measured goodness of fit using the Likelihood Ratio Index (LRI) as a pseudo R^2, defined as:

(2) $1-(Lu/Lr)$,

where Lu and Lr are the unrestricted and restricted log likelihood values, respectively.

Precision or efficiency of the WTP (EFWTP) estimate will be measured by comparing a "standardized" confidence interval around the mean WTP estimate. We standardize by dividing the 95 percent confidence interval by the mean WTP giving a CI/Mean for comparison purposes.

Therefore our specific hypotheses for goodness of fit and for precision are:

(3) H^1_O: $LRI_{LIMATT} = LRI_{SUM} = LRI_{ASUM} = LRI_{STD} = LRI_{YES}$

(4) H^1_A: $LRI_{LIMATT} > LRI_{SUM} > LRI_{ASUM} > LRI_{STD} > LRI_{YES}$

(5) H^2_O: $EFWTP_{LIMATT} = EFWTP_{SUM} = EFWTP_{ASUM} = EFWTP_{STD} = EFWTP_{YES}$

(6) H^2_A: $EFWTP_{LIMATT} > EFWTP_{SUM} > EFWTP_{ASUM} > EFWTP_{STD} > EFWTP_{YES}$

with YES representing YES10, YES910, and YES810 collectively.

Lastly, we model the determinants of respondent uncertainty. We hypothesize that response uncertainty changes with the bid amount and that this effect of the bid amount is influenced by how the respondent answered the WTP question. For respondents answering YES, response to a low bid of $1-3 should be fairly certain while a YES to high values should reflect a high amount of uncertainty. The opposite may be true for those answering NO, as they should become more certain at higher bid amounts.

As suggested by Wang (1997) respondents will become more certain of the YES answer as the bid amount they are asked to pay falls farther below their actual mean WTP and more certain of NO responses as the bid increases above their actual mean WTP. When the bid amount they are asked to pay is close to their actual mean WTP, there is much uncertainty (in Wang's work, these were "Don't Know" responses). Therefore, one additional test of the data is formalized by the model of response uncertainty:

(7) $\text{Certainty} = A_0 - A_1(\text{Bid}) + A_2(\text{Bid}^2) + A_3(\text{Inced}) + A_4(\text{Knowfish})$

where Certainty is the respondent's rating of their certainty in the WTP answer, Bid is the dollar amount they are asked to pay with Bid^2 being the square of Bid, Inced is the cross product term of income and education, and Knowfish is respondent's knowledge about T&E species.

Given this model, our hypothesis is:

(8) $H^3{}_0$: $A_1(\text{Bid}) = A_2(\text{Bid}^2) = A_3(\text{Inced}) = A_4(\text{Knowfish}) = 0$

(9) $H^3{}_A$: $A_1(\text{Bid}) < 0; A_2(\text{Bid}^2) > 0; A_3(\text{Inced}) > 0; A_4(\text{Knowfish}) > 0$

Statistical Estimation of the Logit Model

Since the printed dollar amount varies across the sample of respondents, the voter referendum format requires the analyst to statistically trace out a demand-like relationship between probability of a YES response and the dollar amount using a qualitative response model such as logit or probit (Hanemann 1984). The basic logistic regression model was given in equation (1). From equation (1), Hanemann (1989) provides a formula to calculate the mean or expected value of WTP assuming WTP is greater than or equal to zero. The formula is:

(10) $\text{Mean WTP} = (1/B_1) * \ln(1+e^{B_0})$ where $\text{WTP} \geq 0$,

with B_1 being the coefficient estimate on the bid amount and B_0 is either the estimated constant (if no other independent variables are included) or the grand constant calculated as the sum of the estimated constant plus the product of the other independent variables times their respective means. In this research, B_0 is the grand constant.

As can be seen in equation (10), calculation of mean WTP from a logit model involves the ratio of two random variables. Therefore, this must be recognized when calculating the confidence interval around the mean WTP. The approach of Park et al. (1991) does this using the variance-covariance matrix of the estimated logit equation.

CASE STUDY PROTECTING CRITICAL HABITAT FOR NINE T&E FISH SPECIES IN RIVERS OF THE FOUR CORNER STATES

This case study uses the CVM to quantify public economic value for preserving critical habitat units (CHUs) that are habitat to nine T&E fish species in the four corners region of the U.S. Nine species of fish are listed as threatened or endangered and have critical habitat designated in six rivers of the four corners states as shown in Table 1. The impact of having critical habitat designated is that river flows are affected through instream flow requirements and altering management of hydropower facilities.

CHUs are designated as necessary for survival and recovery of a designated species under Section 7 of the ESA. These areas allow for recovery of these fish species with the goal that they will be removed from listing. The primary habitat components for these fish species are rivers that provide or have the potential to provide life requisites. Criteria for delisting requires a stable or increasing population after 10 years and habitat trends must be stable or increasing over the long-term (U.S. Department of the Interior 1995).

The U.S. Fish and Wildlife Service designated CHUs on 2,456 river miles including segments in major cities such as Grand Junction, Colorado and Albuquerque, New Mexico, plus the Colorado River downstream of Glen Canyon Dam. These include the Colorado River through Glen Canyon National Recreation Area and Grand Canyon National Park. Portions of the Gila River, including stretches through Phoenix, Arizona are also designated.

Survey Design

Prior to designing the actual survey, focus groups were held in Fort Collins, Colorado, Albuquerque, New Mexico and Phoenix, Arizona, leading to revisions based on the suggestions and comments of the participants in these groups. Following the focus groups, the research team developed a complete mail booklet and survey script, used to pre-test a small sample of households throughout the U.S.

Feedback suggested further refinements, a more explicit voting emphasis, ways to reduce repetition and improvement in survey instructions. Responses to the pre-test bid amounts formed the basis (along with on going research on the economic value of the Silvery Minnow by Robert Barrens at University of New Mexico) in establishing the bid amounts in the final survey. The final questionnaire was typeset into a 12 page booklet.

Survey Structure

The first section of the survey allowed the respondents an opportunity to reflect on why they might care about the endangered species and was used for collecting their thoughts on the topic (Cummings et al. 1986). The first set of questions asked about the relative importance of federal lands for providing habitat for endangered species versus using resources for extraction and jobs. A five point Likert scale allowed individuals to agree or disagree with a set of attitude questions to measure how utilitarian they were versus how preservation oriented they were. These responses also provided insight into the responses to the WTP question.

Our CVM survey followed the standard three element design: (a) portrayal of the resource to be valued, (b) description of the particular mechanism to be used to pay for the resource and (c) the question format used to elicit the respondent's dollar amount of WTP. The resource being valued was the 2,456 miles of CHUs described earlier. Survey respondents were provided detailed maps with the CHUs highlighted. Protection involved habitat improvements such as fish passageways as well as bypass releases of water from dams to imitate natural water flows needed by fish. Table 1 shows the listing of fish species by river that was printed in the survey.

Households were told that some State and Federal officials thought the costs of the habitat improvements and the restrictions on hydropower were too costly and proposals for eliminating CHUs had been put forward. Then the description of the particular mechanism to be used to pay for the resource was provided. They were told the current program could be paid for by the establishment of a Four Corners Region Threatened and Endangered Fish Trust Fund. Efforts to raise funds would involve all U.S. taxpayers contributing to this fund. If a majority of households vote in favor, the fund would maintain CHUs for the nine Threatened and Endangered fish species to avoid extinction. This would be accomplished through water releases from Federal dams timed to benefit fish and the purchase of water rights to maintain instream flows. The survey stated that within the next 15 years, three fish species would increase in population to the point they would no longer be listed as a Threatened Species.

However, if a majority of households in the U.S. vote to not approve. then the CHUs shown on the enclosed map would be eliminated. That would mean water diversion activities and maximum power production would occur, reducing the amount of habitat for these nine fish species, and that as a result, biologists estimate that it is very likely that four of the nine fish species will be come extinct in 15 years.

This information was followed by the question format used to elicit the respondent's dollar amount of WTP which asked each household how they would vote, considering the price indicated. This referendum format is recommended by the "blue ribbon panel" on CVM (Arrow et al. 1993). The exact wording on the questionnaire was:

Suppose a proposal to establish a Four Corners Region Threatened and Endangered Fish Trust Fund was on the ballot in the next nationwide election. How would you vote on this proposal?

Remember, by law, the funds could only be used to improve habitat for fish.

1. *If the Four Corners Region Threatened and Endangered Fish Trust Fund was the only issue on the next ballot and it would cost your household $\$$_____ every year, would you vote in favor of it? (Please circle one.)*

YES NO

The dollar amount, which is blank in this example, was filled in with one of 14 amounts ranging from $1 to $350, randomly assigned to survey respondents. The range was picked such that at the low end, anyone that valued preserving the fisheries protection would very likely indicate they would pay $1-3, while almost no one was expected to pay $350 per year.

On the next page of the survey, respondents was asked to determine how certain they were when answering the WTP question. The wording in the survey was as follows:

2. *On a scale of 1 to 10, how certain are you of your answer to the previous question? Please circle the number that best represents your answer if 1=not certain and 10= very certain.*

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

not certain< - > very certain

Sample Frame and Survey Mailing

The questionnaire was sent to a random sample of 800 households in the four corner states of Arizona, Colorado, New Mexico and Utah (with the proportions based on the states relative populations) and an additional 800 households in the rest of the U.S. The sample was provided by Survey Sampling Inc., a company that specializes in providing representative samples and one that has been frequently used by researchers in the past. The overall survey design and mailing procedure follows Dillman's (1978) Total Design Method (first mailing, postcard, second mailing). Each individual was sent a personalized cover letter on university letter head with an original signature. A dollar bill was included with the first mailing as a token of appreciation and to increase the response rate. Both the outgoing and return envelopes had a first class postage stamp affixed to further distinguish the mailing from bulk mail. A second mailing was performed (without the $1 bill) to non-respondents.

Survey Results

We received 718 responses, after deleting undeliverable surveys and deceased, yielded a response rate of 53.9 percent. In the first part of the survey, the respondents were asked about their prior knowledge of three issues related to endangered species, with summary information shown in Table 2. They were asked if they had read or heard about the northern and Mexican spotted owls and the threatened and endangered fish in the Colorado River. Most of the respondents (over 80 percent) indicated prior knowledge of the northern spotted owl. A high proportion of four corner residents had knowledge of the threatened and endangered fish species in the Colorado River (about 74 percent) but only 47% of rest of U.S. residents had knowledge of these species. Another high profile news item in this region was the listing of the Mexican spotted owl and about 55 percent of the four corners residents heard of this species while only 25 percent of the rest of the U.S. responded affirmatively to this question.

Very few respondents indicated membership in environmental organizations with only 12 percent of four corners residents and 14 percent of the rest of the U.S. sample stating they belong to at least one of these organizations. This is an encouraging sign regarding the representativeness of the sample. That is, we did not receive surveys just from those strongly interested in the environment.

Besides the bid amount, independent variables include income, education, and a shift variable representing knowledge of T&E fish, plus proxies for tastes and preferences called PROTECT and PROTJOB, which are explained below. As income and education are highly correlated, these two variables are analyzed as one variable, created by multiplying income and education together and scaled, creating a cross-product called INCED. Knowledge of T&E fish in the four corners region, represented by a variable named KNOWFISH, is the response to a question asking if the respondent was familiar with these fish species.

PROTECT was the sum of the answers on the Likert scale from the questions asking about the desirability of protecting plants and animals. PROTJOB was the sum of the responses related to the rights of business to extract resources and be protected from loss of jobs. As the variable, PROTECT, was the sum of 4 questions and the variable, PROTJOB, was the sum of two questions, PROTECT was divided by two so that the coefficient would compare to the PROTJOB coefficient. Also, because the Likert scale asked the respondent to answer "1" for strongly agree and "5" for strongly disagree, the coefficients to these variables would be intuitively reversed. Therefore, each was multiplied by a negative one to reverse the signs of the coefficients.

Statistical Results

Table 3 provides the coefficients and t-statistics for the logit equations. The coefficients on almost all variables except KNOWFISH are statistically significant. The coefficients for the bid amount are negative, indicating that as bid amount increases, the respondent is less likely to pay this amount. The coefficients for INCED have positive signs, as income and education increases, willingness to pay increase. The coefficients for KNOWFISH are also positive, indicating that those with more knowledge about the species are also willing to pay additional amounts.

PROTECT generated positive signs on these coefficients and are significant showing strong preferences about protecting endangered species and were more willing to bid higher dollar values. Those respondents with high scores on the PROTJOB variable emphasize employment above T&E protection and this results in significant negative signs on the coefficients implying less likelihood of paying to protect the nine T&E fish.

The mean WTP values were calculated using equation 10 with the resulting values shown in table 4 with a range of 95 percent confidence intervals calculated using the variance-covariance matrix (Park et al. 1991). For all the models, the confidence intervals did not include zero, indicating that they all generated significant positive values. The mean WTP for the standard referendum question was $195 per household. This is the value for protecting all nine fish in the seven rivers. In particular, the WTP is to avoid extinction of four fish species and increase the population of three species so they can be delisted. Our values appear to pass an informal scope test by comparing them to the $28.73 value estimated by Barrens, et al. (1996) for just the silvery minnow in the Rio Grande River in New Mexico (one of the nine fish species in our study).

The mean WTP values with the uncertainty treatments are also shown in Table 4. YES10 had the lowest mean of $28, a value much lower than the standard DC estimate, similar to results of Champ et al. The estimated mean WTP for the YES910 model was $52 while the YES810 model generated a mean of $89. As these models include more uncertainty within the YES answer, they move closer to the standard DC model, though for these three models they were significantly different than the mean for the standard DC model. The mean WTP for the ASUM model was $139, $221 for the SUM model and $247 for the LIMATT model. These last three had confidence intervals which overlapped the interval of the standard DC model implying the means were not significantly different.

Results of Hypothesis Tests

Hypothesis #1 suggested that the models using uncertainty information should show improvement in goodness of fit using the LRI. Generally, we found this to be not true. The models which re-code uncertain YES responses to NO (YES10, YES910, and YES810) had slightly smaller LRI estimates as predicted by hypothesis #1. However, we also hypothesized that the next three models that more fully utlize the uncertainty information in the estimation of the logit model should have better goodness of fit, but the opposite was true. The ASUM model's LRI was slightly less than the standard DC model. The other two models had much lower LRIs at .11 and .14 compared to the standard DC value of .24. While Li and Mattsson suggest their approach is an improved structure for CVM, this dataset suggests it has the poorest performance in terms of goodness of fit (.11) of all the approaches.

Table 4 also presents the 95 percent confidence intervals as well as our measure of the precision of estimated WTP called EFWTP, which is the confidence interval divided by the mean. The smaller the value, the greater the efficiency in the estimate of the mean WTP. For our models of uncertainty, the YES910 and YES810 models had an efficiency of .81, close to the standard DC model, with a ratio of .73. Unfortunately, models that more fully incorporated information on uncertainty generally had the largest ratios, not as hypothesized. Again, the Li and Mattsson approach performed the poorest with a 95 percent confidence interval 25 percent larger than the mean WTP. In general, while we reject the of equality of efficiency scores, the alternative hypothesis is rejected as the efficiency was worse. not better with more uncertainty information.

For both hypotheses and for the t-statistics on the variables, the ASUM model performed better than the SUM and LIMATT models. This may imply that, at least for this dataset, scaling uncertainty for the NO respondents may not reflect the nature of the responses. Uncertain NOs may be NOs as shown in Ready et al. Overall, it may be that including the level of respondent uncertainty in the logit model estimation adds more variance than it reduces.

The results of the statistical test of Hypothesis #3 regarding the pattern of respondent uncertainty showed interesting results. First using the complete sample, the signs on all the variables were as predicted, as shown in table 5. However, the level of statistical significance was low for a number of coefficients. For the sub-sample using just the respondents answering YES to the WTP question, the signs again were as predicted with a significant bid variable. For those answering NO, the signs again are as predicted, but only two of the variables had statistically significant coefficients. Nonetheless, the pattern of respondent uncertainty suggested by Wang (1997) does appear to hold. That is, at very low and very high bid amounts respondents are fairly certain of their YES and NO answers, but when the bid amount they are asked to pay is in the mid-range, close to their WTP, they are less certain.

Policy Implications

For the purpose of policy decisions, this study showed there are significant values to protecting these nine T&E fish species. All of the estimates of mean WTP are statistically different from zero, with the lower bound of the most conservative estimate being $20 (Yes10 model). Using the range of all the models, the lowest mean WTP per household is $28 from the Yes10 model. The highest mean is from the Li and Mattsson model, a WTP of $247. The standard dichotomous choice CVM recommended by the NOAA panel yields an estimate of $195 per household. While these values represent a broad range, a resource manager can recognize that if there are close to 100 million households in the U.S., the economic benefit of protecting these habitats is substantial. As noted in the introduction, the Yes10 model has been shown to have some claim to meeting criterion validity, so that the national benefit estimate could be about $1 billion, even assuming non-respondents have zero WTP.

CONCLUSION

This study found WTP for preservation of critical habitat areas for the nine threatened and endangered fish species in the four corners region of the U.S. to be $195 using the standard dichotomous choice model for estimating nonmarket values. This value is substantially greater than the value found by Barrens et al. for one of these nine species, suggesting both estimates pass an informal test of scope or scale. The confidence intervals around our mean WTP did not include zero, implying that this WTP was statistically positive. Efforts at recoding the YES responses similar to those by Champ et al. and Johannesson et al. also resulted in reduced levels of WTP, with confidence intervals not overlaying those of the standard dichotomous choice model. The results of Champ et al. and Johannesson et al. implied that recoding provided for a more realistic estimate of hypothetical WTP.

For our ASUM model which scaled the uncertainty of the YES responses, the WTP was less than the estimate from the standard dichotomous choice model, however, not significantly different. The final two models, which scaled both the YES and NO uncertainty levels, the WTP was greater than the standard dichotomous choice model, but again not statistically different from the standard dichotomous choice model. Past literature suggested that models incorporating uncertainty would be more efficient at estimating the WTP. However, this expectation did not prove to be true, as the pseudo R^2 and the range of the confidence intervals were not effectively improved by the models using the uncertainty information. The degree of uncertainty is

self-assessed by survey respondents and may contain a level of uncertainty itself. Because of this, the variable may be adding more statistical noise than valuable information to the models, lending to less efficient estimates.

While the results of Champ, et al. and Polasky, et al. suggest validity can be improved by a conservative recoding of uncertain responses to NO responses, this calibration approach is not without some costs. In particular, extreme recoding such as the Yes10 model does reduce the explanatory power of the logit WTP equations and reduces the precision of the estimates. Less drastic modeling approaches are proposed to allow for incorporation of uncertainty. Two of these models allows for incorporating uncertainty on both YES and NO responses. However, like the results of Ready, et al., allowing for uncertainty of the NO responses actually increases mean WTP, although in our case it is not a statistically significant increase. While explicitly incorporating uncertainty into modeling of CVM responses appears promising, more research is needed before one can generalize about its net effect.

REFERENCES

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner and H. Schuman. 1993. Report of the NOAA Panel on Contingent Valuation. Federal Register 58(10):4602-4614.

Barrens, R.. P. Ganderton and C. Silva. 1996. Valuing the Protection of Minimum Instream Flows in New Mexico. Journal of Agricultural and Resource Economics, 21(2): 294-309.

Carson et al. 1997. Temporal Reliability on Estimates from Contingent Valuation. Land Economics 73(2): 151-163.

Champ, P., R. Bishop, T. Brown and D. McCollum. 1997. Using Donation Mechanisms to Value Nonuse Benefits from Public Goods. Journal of Environmental Economics and Management 33(2):151-162.

Cummings, R. G., D. S. Brookshire, and W. B. Schulze. 1986. Valuing Public Goods. Rowan & Allenheld, Totowa, New Jersey.

Dillman, D. 1978. Mail and Telephone Surveys. John Wiley and Sons.

Greene, W. 1992. Limdep Version 6.0 User's Manual and Reference Guide. Econometric Software, Inc. Bellport, NY.

Hanemann, M. 1984. Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses. American Journal of Agricultural Economics 66(3): 332-341.

Hanemann, M. 1989. Welfare Evaluations in Contingent Valuation Experiments with Discrete Response Data: Reply. American Journal of Agricultural Economics 71(4): 1057-1061.

Johannesson, M., B. Liljas, and P. Johansson. 1996. An Experimental Comparison of Dichotomous Choice Contingent Valuation Questions and Real Purchase decisions. Department of Economics, Stockholm School of Economics, Stockholm, Sweden.

Just, R., D. Hueth and A. Schmitz. 1982. Applied Welfare Economics and Public Policy. Prentice Hall, NJ.

Kealy, M. J., J. Dovidio and M. Rockel. 1988. Accuracy in Valuation is a Matter of Degree. Land Economics 64:158-170.

Li, C. and L. Mattsson. 1995. Discrete Choice Under Preference Uncertainty: An Improved Structural Model for Contingent Valuation. Journal of Environmental Economics and Management 28:256-269.

Loomis, J. 1989. Test-Retest Reliability of the Contingent Valuation Method: A Comparison of General Population and Visitor Responses. American Journal of Agricultural Economics 71:76-84.

Loomis, J. 1990. Comparative Reliability of the Dichotomous Choice and Open-Ended Contingent Valuation Techniques. Journal of Environmental Economics and Management 18: 78-85.

Loomis, John and Douglas White. 1996. Economic Benefits of Rare and Endangered Species: Summary and Meta Analysis. Ecological Economics 18:197-206.

Magelby, D. 1989. Opinion Formation and Opinion Change in Ballot Proposition Campaigns. In Manipulating Public Opinion, M. Marolis and G. Mauser (eds.). Pacifc Grove, CA: Brooks/Cole Publishers.

Manski, C. 1995. Identification Problems in the Social Sciences. Harvard University Press, Cambridge, MA.

Mitchell, R. and R. Carson. 1989. Using Surveys to Value Public Goods: The Contingent Valuation Method. Resources for the Future, Washington DC.

Park, T., J. Loomis, and M. Creel. 1991. Confidence Intervals for Evaluating Benefits from Dichotomous Choice Contingent Valuation Studies. Land Economics 67:64-73.

Polasky, S., O. Gainutdinova and J. Kerkvliet. 1996. Comparing CV Responses with Voting Behavior: Open Space Survey and Referendum in Corvallis, Oregon. in Herriges, J., compiler, W-133 Benefits and Costs Transfer in Natural Resources Planning, 9th Interim Report. Department of Economics, Iowa State University, Ames, IA.

Ready, R., J. Whitehead and G. Blomquist. 1995. Contingent Valuation When Respondents are Ambivalent. Journal of Environmental Economics and Management 29(2):181-197.

U.S. Department of Interior, Fish and Wildlife Service. 1995. Recovery Plan for the Mexican Spotted Owl (Strix Occidentalis lucida): Vol. I. Albuquerque, NM. 172pp.

U.S. Department of Interior. 1986. Natural Resource Damage Assessments; Final Rule. Federal Register 51. Washington DC.

U.S. District Court of Appeals (for the District of Columbia). State of Ohio v. U.S. Department of Interior. Case No. 86-1575. July 14, 1989.

U.S. Water Resources Council. 1983. Economic and Environmental Principles and Guidelines for Water and Related Land Resources Implementation Studies. Washington DC.

Wang, H. 1997. Treatment of Don't Know Responses in Contingent Valuation Surveys: A Random Valuation Model. Journal of Environmental Economics and Management, 32(2):219-232.

Welsh, M. and R. Bishop. 1993. Multiple Bounded Discrete Choice Models. in W-133 Benefits and Costs Transfers in Natural Resources Planning, 6th interim report, J. Bergstrom, compiler. Department of Agricultural Economics, University of Georgia, Athens, GA.

Measuring the Value of Protecting Ground Water Quality: Results and Methodological Findings

Donald J. Epp, Professor of Agricultural Economics
Willard A. Delavan, Graduate Research Assistant

Department of Agricultural Economics and Rural Sociology
Penn State University

Presented at the Meeting of the Regional Research Project W-133, March 21, 1997, Portland, OR

Abstract

Estimated mean willingness to pay for a hypothetical groundwater protection program in Southeastern Pennsylvania, was $47.16 using an informed open-ended format and $67.85 using a format that employed a dichotomous choice question followed by an open-ended question. Effects of prior information, question wording and framing are analyzed.

Introduction

Of the objections to contingent valuation method (CVM) approaches to estimating the demand for environmental quality, perhaps the most difficult to counter is the susceptibility of willingness to pay bids to differ with variations in the wording of the valuation question. It is a formidable challenge to present a hypothetical choice situation and make it sufficiently realistic that respondents are both able and willing to give well thought out, preference ordered bids within their budget constraints. Developments over the past ten years have allowed researchers to reach agreements on some aspects of contingent scenarios. For example, CVM questionnaire development should include extensive verbal protocols to design a questionnaire that reduces the possibility that respondents will consider something other than what the researcher wants valued. Also, it is generally accepted that dichotomous choice questions may suffer from starting point bias while open-ended questions result in wide bid ranges due to respondents unfamiliarity with assigning a price to things that are not purchased. In this paper we address one such concern--the effect of two different valuation question formats on the estimates of willingness to pay for a hypothetical groundwater protection program in southeastern Pennsylvania.

Literature Review

Mitchell and Carson (1984) divide elicitation techniques into four main types: bidding games, take it or leave it, payment card, and open ended. The first two present an amount and the respondent either refuses or accepts it. The payment card presents several to many amounts from which the respondent chooses one. The last type, the open ended valuation question, allows the respondent to select any amount. Since the hypothetical nature of a contingent scenario requires complex preference searching on the part of respondents the elicitation technique plays an important role in that task. When deciding which elicitation technique to use, researchers must consider both the validity of the data obtained and the cost per data point.

With dichotomous choice it is easier for respondents to give a meaningful value yet the rigorous assumptions required for the estimation techniques are difficult to verify due to questions of the mathematical form of the valuation function. Specifically there is a lack of consensus on how to deal with sample design since WTP is sensitive to changes in both the bid range and intervals between payment amounts (Cooper and Loomis, 1992). A single dichotomous choice question, sometimes called take it or leave it, may be easier for the respondent but yields only one data point per respondent. The iterated approaches force respondents to search their preferences more completely and provide more data points. But, they may be prone to compliance bias--the bid may represent a pressured bid higher (or lower) than what the respondent is truly willing to pay. This method is also prone to starting point bias.

Researchers have offered several explanations for this anchoring. Randall and Brookshire (1978) first suggested that an anchoring or starting point bias stems from an inability to define or perceive the good to be valued on the part of respondents. They conjectured that if the starting bid is significantly different from the respondent's actual WTP, the bidding process may bore them into prematurely offering any bid to end it or to a second possibility that the initial bid, in effect suggests a range of possible final bids. Boyle et al (1985) posit that the initial bid may suggest a reasonable final bid *because* the respondents are unfamiliar with the good and/or the elicitation technique.

To avoid the potential biases of iterative bidding, researchers developed a payment card listing a range of possible bids. This reduced starting point bias of bidding games and the high nonresponse problem of open ended questions.

When developing NOAA's final guidelines for CVM studies it was acknowledged that the dichotomous choice format better replicates an actual referendum whereas the open-ended format is completely unfamiliar. But, the open-ended format was endorsed for use in liability cases by the NOAA panel which noted that open ended questions often produce lower estimates of WTP, albeit with higher variances. A few studies have compared the two elicitation formats. Some split sample studies (Johnson et al., 1990; McFadden, 1993; and Boyle et al., 1993) found dichotomous choice estimates significantly higher than open-ended estimates. Brown et al. (1996) documented eleven comparisons in which the ratio of dichotomous choice to open ended bids ranged from 1.6 to 4.4. Kealy and Turner (1993) showed that although intended to measure the same theoretical construct (WTP) open ended and closed ended question formats did not yield the same values in their study of wilderness preservation. More recently, Ready et al. (1996) showed that dichotomous choice questions generated significantly larger estimates of WTP than open ended questions.

One explanation for the difference in bids derived from the two elicitation methods is starting point bias or anchoring. Few authors have explicitly tested for starting point bias for such a test requires asking the same respondents both questions which, depending on the order the different elicitation questions are placed in, runs the risk of serious question ordering bias. Boyle et al. (1985) found significant starting point bias in contingent valuation applications of the bidding game. Hanley (1989), however, found no evidence of starting point bias in a CVM study of willingness to pay for protecting nitrates from groundwater protection.

Data and Methods

The study is part of a collaborative effort including CVM studies in several states. The studies attempt to estimate the value of policies to protect groundwater from nitrate contamination. The study reported here employed a mailed survey of residents in portions of Lebanon and Lancaster counties in Pennsylvania. Since much of the study area has nitrate levels that exceed the US EPA maximum contaminant level for public drinking water supplies, it was believed that citizens of the area would take interest in a study concerning policies to protect groundwater quality and provide considered responses to the survey.

Prior to administering the survey, the research team used verbal protocols and pretests of questionnaire items to test whether respondents understood what was asked and to obtain a bid range for policies to protect groundwater quality. Results of the preliminary investigations were combined with results from similar efforts in two other states to design the final questionnaire.

This study employed two forms of the valuation question. After a section of the questionnaire that provided information about groundwater and the health hazards of nitrates in drinking water, one form of the valuation question presented a dichotomous choice question with bid levels selected from the pretests. Rather than following the usual double-bounded dichotomous format of following this question with another dichotomous choice question, the study employed an open ended question as the second valuation question. The second form of the valuation question presented information about the average cost per household of local government expenditures for safety related activities, such as fire protection, police services and the construction and maintenance of streets and highways and followed this with an open-ended question about the maximum amount the respondent's household would be willing to pay for a plan to protect groundwater quality. The annual expenditures for these services cover approximately the same range of dollar values as were used in the dichotomous choice initial bids. (The exact wording of the survey questions may be obtained from the authors).

The study includes an innovative tactic to checking for information bias. We wanted to reduce information bias as much as possible for two reasons. First, to insure that responses were from individuals that are at least minimally knowledgeable about groundwater and nitrates. Second, if information bias can be reduced it is easier to identify anchoring bias. A short quiz follows the information section of the questionnaire to verify respondents knowledge of the subject. While this technique does not distinguish between knowledge respondents already possessed and information they acquired through reading the material provided, knowing that respondents understand the good to be valued is essential and should reduce the likelihood of anchoring due to unfamiliarity with the good to be valued.

Another innovation was to ask respondents to evaluate the likelihood that the water in the study area will remain safe to drink over the next 10 years, first if the program described were approved and second, if the program were not approved. The response to each questions was marked on a line representing probabilities ranging from 0% to 100%. The difference in the two ratings was used as an independent variable when analyzing willingness to pay responses.

The survey was conducted by the authors in the summer of 1996 in a manner adapted from Dillman's Total Design Method (1978). Three mailings were sent at the recommended intervals to 1000 households chosen randomly from telephone listings. The sample was split equally between those receiving the two formats. The response rate was 68%. Useable questionnaires were received from 617 respondents -- 284 for the Dichotomous Choice (DOE) format and 333 for the Informed

Open-ended (IOE) format. Over fifty percent of the zero bids were protest bids. These were excluded from the sample for this study, but are being analyzed in a continuation of the study.

Data Analysis

The mean willingness to pay was calculated from the answers to the open-ended question in each format. The mean WTP of respondents receiving the DOE format was \$67.85 with a standard deviation of 90.47. The responses to the IOE format had a mean WTP of \$47.16 with a standard deviation of 60.25. These mean values are statistically significantly different from each other and the remainder of this section presents results of analyses to examine possible reasons for this difference.

The first step was to compare the socioeconomic characteristics of the respondents of the two sub-samples to determine if they are comparable. If the population means of various characteristics differ for the two sub-samples, it may indicate problems with sample selection and will complicate comparisons of the estimated beta coefficients. T-tests of the relevant variables (income, age, concern for own safety relative to drinking water and priority placed upon water protection) show that there is no significant difference between the two samples on any item except the measure of WTP.

Our analysis consists of two comparisons each using censored regression with the open ended responses of the IOE and DOE questions as the dependent variable. The independent variables are the continuous variables age (AGE), difference between the respondent's perception of safety with and without the program (DIFFERENCE), a dummy variable for household income (HIINCOME, over \$50,000 =1), a dummy gender variable (GENDER =1 if male); a dummy variable indicating whether the respondent uses a private well (PRIVATE WELL=1 if the respondent uses a private well for their drinking water source); and a dummy variable that is the interaction of respondent's concern for their own safety as it relates to drinking water (either concerned or very concerned for self=1); and the priority respondents feel government should place on protecting groundwater (high priority or very high priority=1). The interaction variable is called (CONCERN).

The first regression is a censored (Tobit) regression of the open ended sections. The open ended regression results are listed in Table 1. The IOE regression shows that DIFFERENCE, GENDER, HIINCOME, AGE, and CONCERN are significant at the .01 level or better while PRIVATE WELL is not. The signs are all as expected except for GENDER. The positive significant coefficient on the GENDER variable in both sets of regressions using the IOE responses may indicate that the traditional finding of males being less likely to pay for environmental protection depends on the commodity being valued. We had hypothesized that private well owners would bid higher than those on municipal supplies but the null hypothesis, that there is no significant difference between the two groups cannot be rejected. This could be due to a misperception of the good being valued (they may perceive water as a private good) or it could be due to the influence of those who may have had their wells test negative for nitrates.

The third and fourth columns of table 1 show the results for the censored regression of the same model using data from the open ended portion of the DOE format. Note that almost all parameter estimates are insignificant, a possible indication of some overriding factor in respondents' decision calculus. If we include the initial bid or payment amount as an independent variable (DC-BID) then pseudo r-squared increases and the estimated coefficient for that variable is highly significant with the expected sign. This shows that a large degree of variation may be explained by the starting bid. This points to starting point or anchoring bias and that the direction of that bias is positive. Of course the sign of the bias may have been determined by the fact that some of the initial payment amount questions were higher than the median and mean bids. This result could also be related to survey design, since fewer respondents answered no to the dichotomous choice question and offered a second (lower amount) in the second part than answered yes and offered a higher amount.

The second set of regressions used the specification that best fit responses to the DOE format. The income and age variables were modified where INCOME SQUARED replaces the dummy variable INCOME and the dummy variable MATURE (=1 for those age > 55) replaces AGE.

The results (Table 2) show that the even when the IOE data is regressed on the best model of the dichotomous form coefficient estimates for three of the six independent variables remain significant, while DOE is plagued by an anchoring problem. Perhaps the initial payment amount in DOE differs substantially from the respondents' actual WTP so that the yes responses reflect either a satisficing bid or a rapid recalculation and doubt in respondents' preference orderings.

Statistical equivalence of the two models was tested using the following hypothesis

$$H_0^1 : WTP_{DOE}(X\,;\beta) = WTP_{IOE}(X\,;\beta)$$

Comparison of both the models involved using the likelihood ratio test.

$$LR = 2 \bullet [(ll_1 + ll_2) - ll_{pool}] \_ \chi^2(r)$$

Both tests of equivalence were rejected at the five percent level (chi-square stats of 26.9 and 28.5 for the IOE and DOE bases respectively), consistent with the hypothesis that valuation question wording significantly effects respondents' bids; the null hypothesis is that the two forms are not statistically different, see Poe and Welsh (1996) .

An adequate measure of goodness of fit for Tobit applications has not been consistently employed. Veall and Zimmermann (1994) suggest the use of a pseudo R^2 employed by McKelvey and Zavoina (1975) as the best available measure.

The McKelvey and Zavoina R^2 is : $R^2_{MZ} = \dfrac{\sum_{i=1}^{n} (\hat{Y}^*_i - \overline{Y}^*_i)^2}{\sum_{i=1}^{n} (\hat{Y}^*_i - \overline{Y}^*_i)^2 + N \hat{\sigma}^2}$

The pseudo R^2 are significantly different for IOE specifications but are not for the DOE specifications indicating a relative advantage for the IOE specification in terms of goodness of fit. Note the increase in pseudo R^2 with the addition of the dichotomous choice bid amount suggesting some type of anchoring problem to which we now turn.

The model for starting point bias is based on Boyle et al. (1985) and Bishop and Heberlein (1986). When there is no anchoring, the respondent's true WTP is measured by their final bid, □f. Then the indirect utility function yields a defined level of utility:

$$V (P, e'', Y - \beta_f) = V (P, e', Y) = U$$

where P is a vector of prices, e'' is the level of groundwater safety with the program, e' is the initial level of groundwater safety, and Y is the respondent's income. β_f is the Hicksian measure of consumer surplus represented by:

$$\beta_f = \int_{e'}^{e''} h^{-1}(P, e, U)\, de$$

If an anchoring problem exists then :

$$V(P, e'', Y - \beta_f)\_V(P, e', Y) = U$$

and the corresponding Hicksian measure would not equal the final bid β_f because it is now a function of both the initial bid or anchor point and a vector of variables normally expected to affect WTP (income, age, education level, gender,...). The test for starting point bias is straightforward.

$$WTP(X; \beta) = WTP(X; \beta, \beta_s)$$
$$H_0 : \beta_s = 0$$

The t-statistic (6.4) is significant at the .01 level allowing us to reject the null hypothesis, indicating that the requested payment in the dichotomous choice format serves as a significant determinant of respondents' WTP; it is a significant variable in the respondent's utility function.

Conclusions

Questionnaire design can significantly influence willingness to pay responses in a CVM study. In this study we attempted to isolate anchoring bias by reducing other potential sources of bias, such as information about the nature of good to be valued. Our approach was to include basic information about the good, a test to encourage respondent consideration of the information, and carefully designed questions to elicit a valuation response. We controlled for bias that might result from differing nonuse and aesthetic values by examining a good with little or no nonuse value.

The elicitation format that included a dichotomous choice question followed by an open-ended question exhibited significantly higher WTP bids than did the informed open-ended format. Anchoring appears to be the principle factor in the difference between the results of the two elicitation formats. These results suggest that the differing average WTP bids are due in large part to survey design. Thus, it appears that providing an open-ended question as the second part of a double bounded dichotomous choice bid elicitation format does not avoid the anchoring bias associated with the dichotomous choice. Future research should test whether other combinations of elicitation formats yield more consistent estimates. For example, hybrids of polychotomous choice as suggested by Poe (1996) with the informed open-ended follow up may be promising.

References

Bishop, Richard C. and Thomas A. Heberlein. 1986. "Does Contingent Valuation Work?" in *Valuing Public Goods: The Contingent Valuation Method.* Cummings, Ronald G., David S. Brookshire, and William D. Schulze eds. Totowa , New Jersey: Rowman and Allan Held Publishers.

Boyle, Kevin J., Richard C. Bishop and Michael P. Welsh. 1985. Starting Point Bias in Contingent Valuation Bidding Games. *Land Economics.* 61 (1): 321-330.

Boyle, Kevin J., F. Reed Johnson, Daniel W. McCollum, William H. Desvousges, Richard W. Dunford, and Sara P. Hudson. 1993. "Valuing Public Goods: Discrete Versus Continuous Responses." Department of Resource Economics, Working Paper, University of Maine.

Brown, Thomas C., Patricia A. Champ, Richard C. Bishop, and Daniel W. McCollum. 1996. "Which Response Format Reveals the Truth about Donations to a Public Good." *Land Economics.* 72(2):152-166.

Carson, Richard. 1991. *Measuring the Demand for Environmental Quality.* in Braden, John B., and Charles D. Kolstad. eds Elsevier Science Publishers B.V.: Amsterdam.

Cooper, Joseph and John Loomis. 1992. Sensitivity of Willingness-to-Pay to Bid Design in Dichotomous Choice Contingent Valuation Models. *Land Economics.* May 211-224.

Dillman, Don A. 1978. *Mail and Telephone Surveys: The Total Design Method.* New York: Wiley.

Hanley, Nick. 1989. *"Problems in valuing environmental improvements resulting from agricultural policy changes: The case of nitrate pollution"* Discussion paper no. 89/1 Economic Dept, University of Stirling.

Hanley, Nick, and Clive L. Spash. 1993. *"Cost Benefit Analysis and the Environment."* Edward Elgar Publishing Co. Brookfield VT.

Johnson, Rebecca L., N. Stewart Bregenzer, and Bo Shelby. 1990. "Contingent Valuation Question Formats: Dichotomous Choice versus Open-Ended Responses." In *Economic Valuation of Natural Resources*, ed. R.L. Johnson and G.V. Johnson. Boulder, CO: Westview Press.

Kealy, Mary Jo and Robert W. Turner. 1993. "A Test of the Equality of Closed-Ended and Open ended Contingent Valuations." *American Journal of Agricultural Economics.* 75 (May):321-331.

Maddala, G.S. 1992. *Introduction to Econometrics*. New York: MacMillan Publishing Company.

McFadden, D. and G. Leonard. 1993. "Issues in the Contingent Valuation of Environmental Goods: Methodology for Data Collection and Analysis." *In Contingent Valuation: A Critical Assessment*, ed. J. Hausman. Amsterdam: Elsevier.

McKelvey, R. and W. Zavoina. 1975. "A Statistical Model for the Analysis of Ordinal Level Dependent Variables." *J. Mathematical Sociology*. 4:103-20.

Poe, Gregory L., and Michael P. Welsh. 1996. *Uncertainty in Contingent Values and Procedural Variance Across Elicitation Formats: A Multiple Bounded Discrete Choice Approach.* Working Paper 96-03, Cornell University.

Randall, Alan, and Brookshire, David S. 1978. *"Public Policy, Public Goods and Contingent Valuation Mechanisms."* Paper presented at the Western Economic Association Meeting, Honolulu, Hawaii.

Ready, Richard C., Jean C. Buzby, and Dayuan Hu. 1996. Differences Between Continuous and Discrete Contingent Value Estimates. *Land Economics*. 72(3):397-411.

Veall, Michael R. And Klaus F. Zimmermann. 1994. Practitioners Corner Goodness of Fit Measures in the Tobit Model. Oxford Bulletin of Economics and Statistics. 56(4):485-499.

Table 1

Tobit Regressions--IOE Base

	Informed Open-ended (IOE) n = 275			Dichotomous Open-ended Follow Up (DOE) n =225			Dichotomous Open-ended Follow up with DC-BID n=225		
Variable name	[ ]	t-ratio	prob: t > x	[ ]	t-ratio	prob: t> x	[ ]	t-ratio	prob: t> x
CONSTANT	-34.294	-1.114	0.265	-25.183	-0.527	0.598	-66.288	-1.464	0.143
DIFFERENCE	1.990	6.491	0.000	3.008	6.219	0.000	2.976	6.702	0.000
GENDER	52.496	3.142	0.002	-13.389	-0.504	0.614	-18.081	-0.739	0.459
CONCERN	28.328	1.855	0.064	27.373	1.142	0.253	15.034	0.676	0.499
HIINCOME	30.931	2.034	0.042	31.347	1.270	0.204	24.576	1.076	0.281
AGE	-1.1385	-2.262	0.024	-0.221	-0.311	0.756	-0.155	-0.236	0.813
PRIVATE WELL	-11.943	-0.793	0.428	-26.594	-1.137	0.255	-20.610	-0.953	0.340
DC-BID							0.322	4.336	0.000
log-likelihood	-911.7674			-886.0311			-874.7229		
pseudo R-squared	0.2733			0.2446			0.3271		

Table 2

Tobit Regressions--DOE Base

	Dichotomous Choice w/out DC-BID n = 226			Dichotomous Choice with DC-BID n = 226			Informed Open-ended n = 263		
Variable name	□	t-ratio	prob: t > x	□	t-ratio	prob: t > x	□	t-ratio	prob: t > x
Constant	-38.113	-1.411	0.158	-92.477	-3.365	0.000	-77.822	-4.068	0.000
INCOME SQUARED	0.009	2.293	0.02	0.001	2.655	0.008	0.003	1.124	0.260
DIFFERENCE	2.78	5.894	0.000	2.853	6.642	0.000	2.736	7.065	0.000
MATURE	2.614	0.097	0.923	13.523	0.550	0.582	-28.068	-1.343	0.179
PRIVATE WELL	-30.897	-1.403	0.604	-23.783	-1.185	0.236	-16.680	-1.041	0.298
CONCERN	63.535	2.216	0.0276	41.914	1.590	0.112	36.381	1.854	0.064
GENDER	-14.748	0.592	0.554	-20.262	-0.895	0.370	50.006	2.830	0.005
DC-BID				0.348	5.054	0.001			
log-likelihood		-882.44			-870.97			-911.20	
Pseudo R-squared		0.2761			0.3474			0.2713	

NONLINEAR INCOME EFFECTS IN RANDOM UTILITY MODELS*

Joseph A. Herriges and Catherine L. Kling
Department of Economics Iowa State University

*The authors would like to thank Daniel McFadden for his assistance in implementing the GEV sampling routine and Douglass Shaw for his detailed comments and suggestions on an earlier draft of this paper. We would also like to acknowledge the careful reviews of three anonymous referees and the comments and suggestions made by participants of the W-133 Western Regional Project's annual conference. As usual, all remaining errors are our own. This research was funded, in part, by Iowa State University's Agricultural Experiment Station and a grant from the U.S. Environmental Protection Agency. Although the research described in this article has been funded in part by the United States Environmental Protection Agency through R82-3362-010 to the authors, it has not been subjected to the Agency's required peer review policy and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

Abstract

The random utility model (RUM) is commonly used to represent the individual's allocation of trips to a set of recreation areas. Empirical application of the RUM is performed using the conditional logit model. This model provides well-known measures of per-trip consumer surplus, but scaling up these measures to aggregate seasonal or annual values is problematic because the underlying multinomial logit model conditions on the total number of trips to all sites. As a result interest has focused on linking the conditional logit trip allocation model to a model of aggregate demand using a price index derived from the RUM. Using results on the logit representative consumer of Anderson et al., this paper shows that a utility theoretic aggregate price index that is consistent with a logit allocation model model does not exist when the aggregate good is defined as total recreational trips. If aggregate demand is defined as total recreational travel and if the conditions for Hicks composite commodity theorem are satisfied, then it can be shown that trip allocation and total travel demand can be determined in a utility theoretic manner and welfare measures can be derived. The paper presents a conditional indirect utility function which links a site allocation model of logit form to a proper aggregate demand model.

I. Introduction

Constant marginal utility of income is viewed in most microeconomic applications to be a restrictive, special case of the more plausible scenario in which marginal utilities vary with respect to both prices and incomes. Yet in the context of modeling discrete choices made by consumers (e.g., the selection of travel mode or which recreation site to visit), analysts have relied almost exclusively on random utility models that are linear in income, directly imposing a constant marginal utility of income.[1] This assumption is common even in cases where the estimation of welfare measures is the primary goal of the empirical work where nonlinear income effects are likely to be important (See, e.g., Just, Hueth, and Schmitz, (1982)).

The imposition of linear income effects has been accepted in part because of the inconvenience of estimating nonlinear models, but more importantly because of the difficulty of computing welfare estimates under these circumstances.[2] In fact, methods for computing welfare estimates using nested logit models that allow for nonlinear income effects have only recently been devised (McFadden (1995)) and have not previously been implemented using actual data. Unfortunately for the practitioner, the procedures outlined by McFadden are computationally intensive, requiring repeated draws from a random sampler for the generalized extreme value (GEV) distribution and an iterative algorithm to implicitly solve for individual welfare impacts. As an alternative, McFadden derives theoretical bounds on these welfare impacts that are computationally simpler than computing point estimates and which, for some applications, may provide sufficient information for policy makers. These recent developments raise the empirical question as to whether nonlinear income effects are important in practice and worth the additional computational burdens that they entail.

The purpose of this paper is to both investigate the empirical consequences of nonlinear income effects in random utility models (RUMs) and to extend and refine the available methods for obtaining welfare estimates in this context. We begin, in section two, by reviewing the basic theory of welfare measurement in RUMs, including results specific to the standard linear model. The next section then identifies the three alternative approaches to computing welfare measures once nonlinear income effects are permitted and discusses the merits of each. We first review McFadden's algorithm for computing willingness-to-pay in a nested logit model with nonlinear income effects and discuss some technical issues related to his proposed resampling scheme. Second, we discuss the alternative of computing welfare measures based upon a representative consumer. This is the approach employed by Morey, Rowe, and Watson (1993) and Shaw and Ozog (1996). Third, we consider in detail the suggestion by McFadden (1995) that bounds alone be computed on the welfare measures of interest. Specifically, we present a modification to his algorithm that increases its accuracy, provide an empirically tractable method for implementing his bounds when there are nonlinear income effects, and identify scenarios in which the welfare bounds are uninformative.

The empirical portion of this work, beginning with sections four and five, is aimed at carefully comparing and contrasting the three alternative strategies for estimating welfare from RUMs. Data from the 1989 Southern California Sportsfishing Survey are used to estimate models of recreational angling that are nonlinear in both income and other arguments of the indirect utility function. Both Generalized Leontief and Translog functional forms are used in modeling the deterministic portion of the utility function. This use of flexible functional forms to approximate the indirect utility function in a RUM is apparently novel. The results are compared to measures constructed from linear models. In addition, several maintained hypotheses about the underlying error distribution are employed, including the extreme value (EV) and several generalized extreme value (GEV) distributions.

In section six of the paper, the estimated models are used to construct welfare estimates for changes in the price of angling, for changes in angling quality, and for the elimination of entire angling sites (due perhaps to the closure of a fishery). We follow each of the strategies for welfare measurement identified above, obtaining both point estimates for the welfare changes and confidence bounds around these estimates. The final section is used to summarize our findings.

II. The Theory of Welfare Measurement in RUMs and the Linear Model

The basic theory and structure of discrete choice random utility models was developed by McFadden (1973, 1974, 1981), Domencich and McFadden (1975), and Diamond and McFadden (1974) for the purpose of analyzing consumer selections from among a set of discrete alternatives and measuring the welfare implications of changes to the available choice set. Early applications focused on transportation choices (e.g., Domencich and McFadden (1975) and Ben-Akiva and Lerman (1985)), though subsequent studies have used this modeling framework to consider issues in education (Gertler and Glewwe (1990)), housing demand (Börsch-Supan (1987)), and energy conservation (Cameron (1985)). More recently, there has been considerable interest in applying RUMs to recreational choices with the primary purpose of computing the welfare implications of changing environmental quality or the loss of access to a recreation area (due, for example, to an oil spill or other environmental disaster) (e.g., Yen and Adamowicz (1994), Hausman, Leonard, and McFadden (1995) and Morey, Rowe, and Watson (1993)).

In discrete choice models, the utility an individual consumer associates with a particular alternative j ($j = 1,\ldots,J$) is assumed to take the form: $U_j = U(z, \mathbf{q}_j, \varepsilon_j)$, where z is the amount of a numeraire good consumed by the individual, $\mathbf{q}_j$ is a vector of characteristics associated with alternative j, and ε_j denotes heterogeneity in consumer preferences and unobserved factors associated with alternative j.[3] The consumer is assumed to choose that alternative yielding the highest utility subject to meeting his/her budget constraint; i.e., $y = p_j + z$ for the selected alternative, where p_j denotes the price of alternative j. Imposing the budget constraint yields the conditional indirect utility functions:[4]

$$U_j = U(y - p_j, \mathbf{q}_j, \varepsilon_j),\ j = 1,\ldots,J. \tag{1}$$

The consumer's problem is then to select the alternative that yields the highest utility.

Using equation (1), the probability of choosing alternative j can be written:

$$P_j(y, \mathbf{p}, \mathbf{q}, \boldsymbol{\varepsilon}) = \text{Prob}\left[U(y - p_j, \mathbf{q}_j, \varepsilon_j) \geq U(y - p_k, \mathbf{q}_k, \varepsilon_k)\ \forall\ k \neq j\right], \tag{2}$$

where $\mathbf{p} = (p_1,\ldots,p_J)'$ and $\mathbf{q} = (\mathbf{q}_1',\ldots,\mathbf{q}_J')'$. The exact form that these choice probabilities will take depends on the assumed underlying distribution for $\boldsymbol{\varepsilon} = (\varepsilon_1,\ldots,\varepsilon_J)'$. If the ε_j's are i.i.d. variates drawn from an extreme value (EV) distribution, then the familiar multinomial specification results, whereas if ε is drawn from a generalized extreme value (GEV) distribution then the nested logit model results.

As noted above, RUMs are often estimated with the goal of measuring the welfare implications of changing the choice set, either the set of alternatives themselves or characteristics of the available alternatives. The compensating variation (cv) for such changes can be implicitly defined by:[5]

$$\underset{j\in \mathbf{J}^0}{Max}\, U\left(y-p_j^0,\mathbf{q}_j^0,\varepsilon_j\right)=\underset{j\in \mathbf{J}^1}{Max}\, U\left(y-p_j^1-cv,\mathbf{q}_j^1,\varepsilon_j\right) \tag{3}$$

where J denotes the choice set and the superscripts "0" and "1" are used respectively to distinguish the original versus new conditions associated with the choice set. The resulting compensating variation is a random variable with the general form

$$cv=cv\left(y,\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,\varepsilon\right). \tag{4}$$

It is the expected value of this random variable that is typically of interest for policy purposes.[6] Unfortunately, there is no general closed-form solution for E(cv), since cv can depend upon the ε_j's in a nonlinear fashion. The standard approach in the literature is to resolve this problem by making the following set of assumptions:

- *A.1 Additive disturbances; i.e.,* $U_j=V\left(y-p_j,\mathbf{q}_j\right)+\varepsilon_j$.
- *A.2 GEV disturbances.*
- *A.3 Constant Marginal Utility of Income; i.e.,*

  $V(y-p_j,\mathbf{q}_j)=\alpha\left(y-p_j\right)+f\left(\mathbf{q}_j\right).$

It can be shown that under these conditions a closed form solution exists for E(cv), one which is independent of income (See, e.g., Hanemann (1982), Small and Rosen (1981), Morey (1994), and McFadden (1995)). While the first two assumptions may be of concern in some applications, the focus in this paper is on relaxing assumption A.3.

III. Relaxing the Linearity Assumption: Implications for Welfare Measurement

The difficulties with relaxing the assumption of linear income effects appear in the computation of welfare measures. One can no longer rely on a closed-form solution to compute welfare. There are currently three alternative approaches from which to choose if one wishes to allow nonlinear income effects.[7] The first alternative requires resampling from the underlying error distribution and employing a numerical algorithm to solve for the implicitly defined compensation variation. The second is to adopt a representative consumer strategy and compute welfare for this consumer. The third approach is to employ McFadden's bounds on the welfare estimates as applied to models with nonlinear income effects. We discuss each option in turn below.

A. Alternative 1: Simulation

The compensating variation defined in equation (3) is an implicit function of the characteristics of the choice set and distribution of preferences in the population, as captured by the functional form of U(·) and distributions of both $(y, \mathbf{p}, \mathbf{q})$ and the ε_j's. Suppose U(·) has been specified to be nonlinear in income and econometric estimates of the parameters have been obtained for a given data set. One approach to computing an estimate of E(cv) is to begin with the simulation procedure suggested in McFadden (1995). The procedure is best understood as a series of steps, conducted first for each observation in the sample:

- *Step 1: At iteration t (t = 1,...,T), a pseudo-random number generator is used to draw the vector $\hat{\varepsilon}^t$ from the estimated distribution of* $\boldsymbol{\varepsilon}$.

- *Step 2: A numerical routine is then used to search iteratively for the* cv^t *implicitly defined by:*[8]

$$\underset{j\in \mathbf{J}^0}{Max}\, U(y - p_j^0, \mathbf{q}_j^0, \hat{\varepsilon}_j^t) = \underset{j\in \mathbf{J}^1}{Max}\, U(y - p_j^1 - cv^t, \mathbf{q}_j^1, \hat{\varepsilon}_j^t). \tag{5}$$

- *Step 3: The mean of the* cv^t *over the T iterations provides a consistent estimate of* $E(cv|y, \mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1)$; *i.e., the mean value of cv for individuals with the set of observed characteristics* $(y, \mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1)$. *The resulting collection of* cv^t*'s likewise provides a simulated distribution of cv for individuals with the same set of observed characteristics.*

If the sample available to the analyst is representative of the target population, these three steps can be repeated for each observation and averaged to obtain an estimate of E(*cv*) for the population. Otherwise, a weighted average may be needed to corrected for differences between the sample and target populations.

This procedure, while conceptually simple, requires the ability to resample from the assumed error distribution used to estimate the model. In this regard, two appealing choices for the distribution of ε are the extreme value and multivariate normal distributions, as pseudo-random number generators are easy to devise in these cases. However, the extreme value distribution yields the multinomial logit specification, which is known to suffer from the much discussed and maligned independence of irrelevant alternatives (IIA) assumption. If one is attempting to generalize the RUM by incorporating nonlinear income effects, it is not likely to be desirable to impose such a restrictive assumption on the disturbance terms. The multivariate normal probit (MNP) model, while certainly less restrictive than the multinomial logit model, is problematic for a different reason. Although recent advances in econometrics suggest that MNP may be feasible to estimate (e.g., McFadden (1989) and Börsch-Supan and Hajivassiliou (1993)), the computational burdens of obtaining parameter estimates for such models remains substantial.

The most common distributional assumption employed with RUMs is that the errors, ε, are drawn from a GEV distribution, resulting in the nested logit model. This specification yields choice probability equations that are easy to construct, thus simplifying estimation, without imposing the IIA assumption which haunts the multinomial logit model. However, approximating a sample from a GEV distribution is not a trivial exercise. In fact, only recently has McFadden (1995) developed a Monte Carlo Markov Chain method for constructing a sequence of random vectors $\hat{\varepsilon}^t$ whose empirical distribution asymptotically approximates a GEV cumulative distribution. The approach to constructing an estimate of E(*cv*) is the same as above, except that Step 1 is replaced with the following GEV sampler routine:

- *Step 1A: At iteration t(t = 1,...,T), a pseudo-random number generator is used to draw J+1 independent (0,1) uniform random variables,* ζ_j^t $(j = 1,\ldots,J)$ *and* η^t. *J extreme value random variates are then formed using the transformation* $\tilde{\varepsilon}_j^t = -\log(-\log(\zeta_j^t))$. *Finally, the following Markov chain is used to construct:*

$$\hat{\varepsilon}^t = \begin{cases} \tilde{\varepsilon}^t & \text{if } \eta^t \leq \dfrac{f(\tilde{\varepsilon}^t)/g(\tilde{\varepsilon}^t)}{f(\hat{\varepsilon}^{t-1})/g(\hat{\varepsilon}^{t-1})} \\ \hat{\varepsilon}^{t-1} & \text{otherwise} \end{cases} \tag{6}$$

where f(·) and g(·) denote the GEV and EV probability density functions, respectively.

The right-hand side of the inequality term in equation (6) can be interpreted loosely by noting that *f(·)/g(·)* corresponds to the weights used in importance sampling (see, e.g., Geweke (1989)). Thus, the Markov chain replaces an earlier draw if the new draw has greater weight than the previous observation in the chain. McFadden (1995) proves that mean compensating variation computed using Steps 1.A. B. and C converges almost surely to $E(cv)$ as $T \rightarrow \infty$.

There are several potential difficulties associated with the simulation estimator outlined above. First. the procedure is computationally intensive. As McFadden (1995) demonstrates in a Monte Carlo experiment. the number of iterations (T) required to achieve a given level of precision increases substantially as the GEV model departs from the EV distribution. In his experiment. the number of iterations required to obtain a five percent root mean squared error ranges from 755 (when ε is EV) to nearly 19.000 as the dissimilarity coefficient becomes 0.1.[9] The computational burden is all the more severe when it is recognized that the parameters underlying these cv calculations are themselves estimates. If confidence bounds on $E(cv)$ are to be constructed recognizing the uncertainty of these estimates. the three-step simulation procedure will need to be repeated for a series of draws from the distribution of the estimated parameters.[10]

Finally. Step 2 of the simulation process assumes that cv^{*} exists which implicitly solves equation (5). This need not be the case when a model with nonlinear income effects is used to approximate underlying preferences. The problem is akin to the difficulties found in continuous demand systems, when estimated models yield preferences that are locally consistent with utility theory, but fail to have well-behaved global properties. In the current problem, while estimated nonlinear models may yield a positive marginal utility of income at the mean of the sample, the marginal utility of income can become negative at extremes of the sample or when substantial price or quality changes are considered. In these cases, there may not exist a cv' which solves equation (5).

B. Alternative 2: A Representative Consumer Approach

A second approach is to approximate $E(cv)$ by computing the income compensation required to equate *expected* utility before and after a given price and/or quality change. This is the approach suggested and implemented by Morey, Rowe and Watson (1993). Formally, this corresponds to calculating the $\overline{cv}$ implicitly defined by:

$$E\left[\underset{j \in \mathbf{J}}{Max}\, U\left(y - p_j^0, \mathbf{q}_j^0, \varepsilon_j\right)\right] = E\left[\underset{j \in \mathbf{J}}{Max}\, U\left(y - p_j^1 - \overline{cv}, \mathbf{q}_j^1, \varepsilon_j\right)\right]. \quad (7)$$

Under this alternative, the expected utility function is interpreted as the utility function of a representative consumer. When preferences satisfy assumptions A.1 and A.2 of the previous section (i.e., ε enters preferences additively and is assumed to be drawn from a GEV distribution), the expected utilities on the left- and right-hand sides of equation (7) are closed form functions of the $V\left(y - p_j^0, \mathbf{q}_j^0\right)$'s and $V\left(y - p_j^1 - \overline{cv}, \mathbf{q}_j^1\right)$'s, respectively.[11] An iterative procedure can then be employed to solve this implicit equation without the resampling step required under the simulation approach. The appeal of this alternative is that it is simple to implement, while still allowing the analyst to relax the constant marginal utility of income assumption (i.e., A.3). However, McFadden (1995) notes that $\overline{cv}$ will generally be a biased estimator of mean compensating variation and, in his Monte Carlo results, finds that the percentage bias from using this approach increases as the size of the welfare change increases.[12] We include estimates based upon this approach in our empirical section below in order to investigate the extent of the bias in an applied setting.[13]

C. Alternative 3: Theoretical Bounds

Recognizing the computational difficulty of the GEV simulation approach, McFadden (1995) suggests that it may be easier to bound the welfare impacts of a policy change and that, for some applications, these bounds may provide sufficient information to decision makers. Towards this end, he proposes theoretical bounds on E(cv). In particular, let cv_{jk} denote the income reduction required to equate the utility from consuming alternative "j" before the quality/price change with the utility from consuming alternative "k" after the change, i.e., cv_{jk} is implicitly defined by:

$$U\left(y-p_j^0,\mathbf{q}_j^0,\varepsilon_j\right)=U\left(y-p_k^1-cv_{jk},\mathbf{q}_k^1,\varepsilon_k\right)\ \forall j,k. \tag{8}$$

The cv_{jk}'s can be viewed as conditional compensating variations, as they are defined as conditional on the event (B^{jk}) that the individual selects alternative j prior to the attribute changes and selects alternative k after these changes and compensation cv. McFadden demonstrates that, given the event B^{jk}, these conditional compensating variations bound the true cv, with[14]

$$cv_{jj}\le cv\left(y,\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,\varepsilon\right)\le cv_{kk}. \tag{9}$$

Taking expectations of this inequality yields McFadden's theoretical bounds on E(cv), with:

$$\sum_{j\in J^0}P_j^0\left(y,\mathbf{p}^0,\mathbf{q}^0\right)cv_{jj}\le E(cv)\le\sum_{k\in J^1}P_k^1\left(y-cv,\mathbf{p}^1,\mathbf{q}^1\right)cv_{kk} \tag{10}$$

where

$$P_j^0(y,\mathbf{p}^0,\mathbf{q}^0) = \text{Prob}\left[U(y-p_j^0,\mathbf{q}_j^0,\varepsilon_j) \geq U(y-p_k^0,\mathbf{q}_k^0,\varepsilon_k)\ \forall\ k \neq j\right] \quad (11)$$

denotes the choice probabilities prior to the attribute changes and

$$P_j^1(y-cv,\mathbf{p}^1,\mathbf{q}^1) = \text{Prob}\left[U(y-p_j^1-cv,\mathbf{q}_j^1,\varepsilon_j) \geq U(y-p_k^1-cv,\mathbf{q}_k^1,\varepsilon_k)\ \forall\ k \neq j\right] (12)$$

denotes the choice probabilities after the attribute changes and compensation cv. The key to employing the welfare bounds in equation (10) is to note that, when the error terms are assumed to enter the utility function in an additive manner (as is typically the case), then the cv_{ij}'s are independent of the error distribution and need not be simulated. Given estimates of the choice probabilities in equations (11) and (12), the computational burden of simulating GEV errors can then be avoided entirely.

There are a number of issues that arise in constructing the theoretical bounds detailed in equation (10). First, while the initial choice probabilities (i.e., $P_j^0(y,\mathbf{p}^0,\mathbf{q}^0)$) follow directly from the estimated model, the new choice probabilities $P_j^1(y-cv,\mathbf{p}^1,\mathbf{q}^1)$ depend upon the unknown compensating variation, cv.[15] Thus, the upper theoretical bound in equation (10) cannot be directly computed. One approach would be to approximate cv using a linear model (in which case cv has a closed form solution) and to use this approximation in computing the upper bound choice probabilities. However, the resulting bounds are no longer guaranteed to truly bound E(cv). Alternatively, the theoretical bounds can be simulated, just as E(cv) can be simulated. In particular, a consistent estimate of $P_j^1(y-cv,\mathbf{p}^1,\mathbf{q}^1)$ can be obtained using McFadden's GEV simulator, which provides a consistent estimator of any real-valued function that is integrable with respect to the distribution function of ε. In this case, that real-valued function is an indicator function for the selected alternative, given the new alternative characteristics and the implicitly solved for cv. Of course, in practice, practitioners would not bother with the theoretical bounds once they had available point estimates of the compensating variation itself.

A third approach to constructing the theoretical bounds is to note that equation (9) implies that

$$cv^L \equiv \underset{i}{Min}\, cv_{ii} \leq cv(y, \mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, \varepsilon) \leq cv^H \equiv \underset{i}{Max}\, cv_{ij}. \tag{13}$$

Using this result, we can bound the new choice probabilities, since

$$\begin{aligned} P_j^1(y - cv, \mathbf{p}^1, \mathbf{q}^1) &= \text{Prob}\left[U(y - p_j^1 - cv, \mathbf{q}_j^1, \varepsilon_j) \geq U(y - p_k^1 - cv, \mathbf{q}_k^1, \varepsilon_k)\ \forall\ k \neq j\right] \\ &\leq \text{Prob}\left[U(y - p_j^1 - cv^L, \mathbf{q}_j^1, \varepsilon_j) \geq U(y - p_k^1 - cv^H, \mathbf{q}_k^1, \varepsilon_k)\ \forall\ k \neq j\right] \\ &\equiv \widetilde{P}_j^1(y, cv^L, cv^H, \mathbf{p}^1, \mathbf{q}^1). \end{aligned} \tag{14}$$

Notice that while $\sum_{j \in J^1} \widetilde{P}_j^1 \geq 1$, it need not sum exactly to unity. Substituting the results of equation (14) into equation (10) yields the following computable bounds on E(cv):[16]

$$\sum_{j \in J^0} P_j^0(y, \mathbf{p}^0, \mathbf{q}^0) cv_{jj} \leq E(cv) \leq \sum_{k \in J^1} \widetilde{P}_k^1(y, cv^L, cv^H, \mathbf{p}^1, \mathbf{q}^1) cv_{kk}. \tag{15}$$

A second issue in the computation of the theoretical bounds is how best to estimate the cv_{jj} terms. McFadden suggests that the equivalent of a linear approximation be employed. This entails computing the difference in utility before and after the quality/price change and dividing by an intermediate value of the marginal utility of income over this change. The accuracy of this approximation is an empirical question, but will undoubtedly decrease as the size of the welfare change increases. However, an exact calculation of the cv_j can be recovered by applying a standard numerical routine (such as numerical bisection) to equation (8).

A third issue regarding these bounds is the type of welfare changes to which they can be meaningfully applied. McFadden writes the bounds in terms of quality changes, corresponding to changes in the level of the **q**'s. We have generalized these in our formulations to consider changes in the prices as well. This is a minor extension, in and of itself, except that it highlights a case when the bounds will be uninformative. Specifically, if the analyst is interested in computing a welfare change associated with the elimination of one or more alternatives (due, perhaps, to toxic contamination of several fishing sites or a large oil spill affecting recreation areas) application of the bounds can be equivalent to considering a price change from its current level to an infinite price for the lost alternatives. For these alternatives, however, cv_{jj} is negative and infinite, since there is no finite level of compensation that can make the consumer as well off with an infinite price compared to the initial finite price, if they are unable to switch to another alternative. Thus, the (absolute value of) the lower bound is infinite. Likewise, the upper bound is zero since it depends on only the alternatives that remain after the site closings and there is no change in the remaining alternatives' prices and/or qualities. In fact, recovery of welfare estimates associated with the entire elimination of one or more alternatives is a common goal of empirical analysis. In these instances, the theoretical bounds will provide no information to policy makers.

A final point concerning the empirical relevance of theoretical bounds in equations (10) and (14) is the recognition that these bounds will need to be computed using parameter estimates, and, as such, will themselves be random variables. Confidence intervals on the theoretical bounds can be computed using simulation techniques. Thus, even if the point estimates of the bounds are fairly tight, the width of the bounds when their statistical imprecision is accounted for may be wider than an analyst is comfortable with.

IV. Data

The data used in our application were drawn from the Southern California Sportfishing Recreation Survey conducted in 1989. A complete description of the data can be found in Thomson and Crooke (1991) and Kling and Thomson (1995). Random telephone interviewing was conducted in Southern California to identify recreational anglers. Those so identified were requested to complete a follow-up mail questionnaire. Respondents provided a variety of information about their angling experiences including extensive information on their most recent saltwater fishing trip. This data included the month of their fishing trip, the species they targeted, the time it took to travel to and from the fishing site, the travel distance, and other expenditures associated with the trip. In addition, they reported whether they fished from the beach, a pier, a private boat, or a charter boat. These four alternatives constitute the possible modes of fishing from which anglers choose in our empirical models.

Respondents also reported their annual income and their household's zipcode. The zipcode data were used to compute roundtrip travel costs to their most recently visited fishing site. An opportunity cost of travel time (based on their reported wage rate) and any boat fees or fuel costs were added to these roundtrip costs to construct the final variable for the price of fishing.

Since the anticipated success of fishing is likely to be an important determinant of the decision to engage in angling as well as the choice of which mode of fishing to select, we include catch rates as an explanatory variable. Specifically, exogenous data on catch rates were provided by the Marine Recreational Fishery Statistics Survey which is sponsored annually by the National Marine Fisheries Service. These catch rates are defined on a per hour fished basis for each major species by fishing mode. In the mail survey, anglers were questioned as to their targeted species. A catch rate variable was then constructed by summing the per hour catch rates associated with each angler's target species. Since these data were collected independently from the mail survey, the catch rate associated with each mode is exogenous to the angler. A total 1182 observations with complete data on income, prices, and catch rates were available for use in our analysis.

V. Model Specification

Our application focuses on modeling the mode choice (i.e., beach, pier, private boat, or charter boat) of recreational saltwater anglers. The model specification involves assumptions regarding the functional form of the indirect utility and the distribution of preferences in the population. We begin by using the standard assumption in the literature (*A.1* above) that the error terms enter the indirect utility function additively; i.e.,

$$U_j = V\left(y - p_j, \mathbf{q}_j\right) + \varepsilon_j \tag{16}$$

where y is now defined as monthly income.

Three alternative functional forms are considered for the deterministic portion of the indirect utility function $V\left(y - p_j, \mathbf{q}_j\right)$. To provide a basis of comparison, we begin by estimating the parameters of a simple linear indirect utility functional form. We also the remaining quartiles, however. the linear approximation yields welfare estimates that are 17 to 24% smaller than the corresponding GEV sampler estimates.

VII. Conclusions

This study has investigated the importance of nonlinear income effects in random utility models, with particular attention to welfare measurement. In addition to specifying a nonlinear structure for the deterministic portion of consumer preferences, using Generalized Leontief and Translog models to provide flexible approximations to any nonlinear utility function, three distinct errors structures were considered. The resulting models were used to study mode choice among California anglers and to compare and contrast the available approaches for computing (or approximately) welfare changes when nonlinear income effects exist. These approaches include a re-sampling scheme based upon McFadden's (1995) GEV sampler, a linear model, a representative consumer approach, and the computation of bounds on the welfare changes of interest. The approaches trade-off computational ease for potential bias in the resulting welfare measures or uncertainty regarding their exact values.

Our analysis of Californian sportfishing represents, to our knowledge, the first application of McFadden's GEV sampler. Several key empirical results emerge. First, our findings (highlighted in Table 2) suggest that, in this application, there are more differences in the point estimates of welfare due to changes in assumed error distribution (e.g., multinomial logit versus nested logit) than due to the introduction of nonlinear income effects. Second, the consistent welfare estimates provided by the GEV sampler are not substantially different from the simpler linear and representative consumer approximations, particularly when the stochastic nature of the underlying parameter estimates is considered. Finally, while the computable bounds are both readily constructed and allow for nonlinear income effects, they do not provide tight

Another important question is whether allowing for uncertain responses ends up actually *encouraging* uncertain responses. i.e., respondents might search their preferences less thoroughly than they would have in the absence of such uncertain response category.

Carson et al. (1995) administer two separate sample of respondents survey instruments that are identical in all respects but the response categories of the vote question: Only one of the two variants of the survey instrument explicitly contains the "would not vote" option. Carson et al. find that when the "would not vote" answer category is explicitly offered to respondents, subjects choose this option more frequently (about 18 percent of the times) than they would spontaneously mention in response to a standard dichotomous choice payment question (8 percent of the times). However, the split between yes and no votes at each level of the bid is not statistically different between the two samples.

3. Statistical Models of Responses

A. The Welsh and Bishop interpretation

Welsh and Bishop (1993) assume that a single, unobserved WTP amount drives all of a subject's responses to the multiple bound payment questions, and develop a statistical framework involving interval data. Specifically, the information coming from all of the responses is collapsed into a single, and relatively tight, interval around the respondent's unobserved WTP value.

To illustrate, consider a person who checks "probably yes" at $5, but "not sure" at $10. One way to interpret this response is to treat the respondent as being willing to pay estimate Generalized Leontief (GL) and Translog (TL) models. Thus, the three specifications considered for the deterministic portion of the indirect utility function are:

- *Linear:*

$$V_j\left(y-p_j,q_j\right)=\beta_{11}\left(y-p_j\right)+\beta_{22}\left(q_j\right) \tag{17}$$

- *Generalized Leontief (GL):*

$$V_j\left(y-p_j,q_j\right)=\beta_{10}(y-p_j)^{1/2}+\beta_{20}q_j^{1/2}+\beta_{11}(y-p_j)+\beta_{22}q_j+\beta_{12}(y-p_j)^{1/2}q_j^{1/2} \tag{18}$$

- Translog

$$\begin{aligned}V_j\left(y-p_j,q_j\right)&=\beta_{10}\ln(y-p_j)+\beta_{20}\ln(q_j)\\&+\beta_{11}\ln(y-p_j)^2+\beta_{22}\ln(q_j)^2+\beta_{12}\ln(y-p_j)\ln(q_j)\end{aligned} \tag{19}$$

where q_j denotes the catch rate at site j. Notice that the linear specification represents a constrained version of the Generalized Leontief model, with $\beta_{10}=\beta_{20}=\beta_{12}=0$.

In addition to identifying the functional form for $V(\cdot)$, the model specification requires distributional assumptions regarding ε. We estimate each model under three assumptions for the distribution of preferences (captured by ε): an extreme value distribution and two GEV distributions corresponding to two different correlation patterns among the alternatives. The extreme value assumption yields the multinomial logit model, whereas the GEV assumptions yield alternative nested logit models, with different nesting structures. The first GEV distribution groups pier, beach, and private boat into a single nest (assuming greater substitution possibilities among these three alternatives than between any one of these and charter boating). This is referred to as the "charter" model. A second GEV distribution is investigated wherein pier, beach, and charter boat enter a single nest and private boat is in its own nest. This is referred to as the "private" model. The tree structures typically presented for nested logit models corresponding to these two correlation patterns, along with the MNL model, are provided in Figure 1.[17]

VI. Results

A. Parameter Estimates

Table 1 contains the parameter estimates from the linear, Generalized Leontief, and Translog functional forms. Coefficient estimates for each these models are reported using the three nesting structures: i.e., the multinomial logit model (MNL), the nested logit charter model, and the nested logit private model. While our primary interest is with the welfare predictions implied by each of these models, several useful insights emerge from Table 1.

Focusing first on the coefficients associated with the deterministic portion of the indirect utility function (i.e., the β_{ij}'s), we see that most of these parameter estimates differ significantly from zero using either a 1 or 5% significance level. While interpreting the β_{ij}'s directly in the nonlinear models is difficult, β_{11} and β_{22} have natural interpretations for the linear models. The coefficient β_{11} corresponds to marginal utility of income and, as expected, is estimated to be positive, ranging between 0.01 and 0.02. Similarly, β_{22} indicates the marginal utility of catch rate (as a quality attribute of fishing mode) and is also estimated to be positive, ranging from 0.41 to 0.95. While these marginal utilities are nonlinear functions in the GL and TL models, their estimates at the sample means were found to correspond closely to those predicted by the linear specification.

The other parameter estimate presented in Table 1 is the dissimilarity coefficient θ. The dissimilarity coefficient indicates the degree of correlation among alternatives within a nest of the assumed nesting structure. A well known condition for consistency of a RUM model with stochastic utility maximization is that θ lie within the unit interval (Daly and Zachary (1979) and McFadden (1978)). When $\theta = 1$, the alternatives are uncorrelated and the multinomial logit specification results. On the other hand, as θ declines towards zero, alternatives within a nest become increasingly closer substitutes. Thus, one test of the multinomial logit specification (and the implied independence of irrelevant alternatives assumption) is whether the parameter θ differs significantly from one. Clearly, the multinomial logit specification is rejected in this application, since θ is statistically different from one using a 1% level for each of the nested logit models. Likelihood ratio tests of this restriction yield the same conclusion. Choosing between the charter and private models is more difficult, since one is not nested in the other. However, since both models have the same number of parameters, application of Pollak and Wales' (1991) likelihood dominance criterion suggests choosing the charter model as the best representation of preferences since (for each functional form specification) it yields a log-likelihood value above the value obtained for the private model.

Finally, the results in Table 1 provide evidence on the statistical validity of the linear model typically employed in the literature. In comparing the flexible forms to the linear model, it is most direct to compare the Generalized Leontief and linear models, since the linear model is nested within the Generalized Leontief. For all three error structures, the linear model is rejected as a restriction on the Generalized Leontief specification using a likelihood ratio test statistic and a 5% significance level.[18] Thus, in general, we find that the more complex GL (and TL) model using either of the nested logit error structures provides a *statistically* better fit of fishing mode choice when compared to the linear multinomial logit model. The question from a policy perspective, however, is whether the more complex models yield *substantially* different welfare predictions, differences that are worth the increased cost of computing welfare impacts in these models. Towards this end, we turn now to a comparison of the welfare predictions using each model specification.

B. Welfare Estimates Using the GEV Sampler

Since the primary purpose of estimating RUMs for recreational angling is to compute welfare measures, we choose three different changes for which to compute welfare impacts. First, we estimate the compensating variation associated with a doubling of the price of each alternative fishing mode. Second, we consider the compensating variation associated with the doubling of the catch rate at all sites and, third, we estimate the compensating variation associated with eliminating two of the modes: pier and beach. The latter change is also a price change, namely one which changes the price of two of the models from their current finite levels to infinity.

Table 2 provides estimates of the mean compensating variation associated with the three changes. The per trip welfare estimates reported in Table 2 were computed using McFadden's GEV sampler and a search algorithm to solve for the implicitly defined *cv* in equation (3). Welfare impacts for each observation in the sample were constructed by averaging estimated *cv*'s computed using T = 1000 iterations.[19] These individual welfare impacts were then averaged over the 1182 observations in the sample.

Several points from Table 2 are worth noting. First, the estimated welfare effects of doubling the price of each mode (Table 2a) is relatively insensitive to the choice of functional form and nesting structure. The welfare loss estimates range only from -\$47.53 to -\$49.74 (the negative indicates that a reduction in welfare occurs). Interestingly, the average price of the four fishing modes, weighted by the original choice probabilities, is just under \$52. The proximity of this average price to the estimated welfare changes in Table 2a suggest that the uniform doubling of prices leads to few mode choice changes, in which case the uniformity of the welfare estimates is not surprising. If no mode changes were to occur, the appropriate compensating variation would simply be this average price.

Turning to Table 2b, we find that quite a different result emerges when a doubling of catch rates is considered for each of the modes. Substantial disparities emerge in the welfare estimates, varying both by the functional form used for the indirect utility function and by the assumed error structure. These estimates ranges from a low of \$7.95 in the case of the Translog charter model to a high of \$26.79 when the linear private model is used, more than a three-fold increase in the estimated E(*cv*). It is worth noting, however, that there is at least as much variability in the welfare predictions due to the choice of nesting structure as there is in form of the indirect utility function.

Finally, in the case of the loss of the shore modes (Table 2c), there is very little difference among the alternative functional forms (reading across the rows of the table), but, again, notable differences among the error structures. For example, given the Generalized Leontief functional form, the welfare impact from losing both shore modes ranges from -\$21.79 to -\$35.24 over the alternative nesting specifications. In contrast, given the charter nesting structure, the welfare loss varies by little more than five percent over the alternative functional forms, from -\$21.79 to -\$22.91.

C. Alternative Welfare Predictions

The welfare estimates for the nonlinear models presented in Table 2 were constructed using the GEV sampler devised by McFadden and, hence, provide consistent estimates of the E(*cv*) associated with each policy scenario, given the selected functional form and nesting structure. However, these computations are quite costly; both in terms of the time required for an analyst to code the algorithms and in terms of computer time.[20] In contrast, the representative consumer approach suggested by Morey, Rowe, and Watson (1993) and the computable bounds in equation (14), based on McFadden's (1995) theoretical bounds, are much easier to obtain. Likewise, as noted earlier in the paper, the linear model avoids the simulation problem entirely, yielding a closed form equation for E(*cv*). Given the ease with which the linear model is estimated and the corresponding ease with which welfare measures (and standard errors) are computed, it may also be reasonable to treat the estimates from the linear model as yet another "second-best" approach to welfare measurement. An interesting empirical question is whether these simpler approaches yield substantially different welfare predictions from those obtained in Table 2. Furthermore, these simpler approaches may be deemed even more palatable when the statistical precision of the point estimates are considered. Thus, if confidence intervals about the point estimates in Table 2 typically encompass the linear or representative consumer approximations, it may be reasonable to compute the simpler measures.

In order to shed light on this issue, Figure 2 provides a comparison of five alternative estimators of the welfare loss due to a doubling of the catch rate.[21] The first alternative is the GEV sampler's estimate of $16.95, indicated by the asterisk in the top bar of the graph. The shade bar around the asterisk represents a 95% confidence bound around the GEV estimate, reflecting the fact that the underlying parameters used in constructing the welfare predictions are themselves random variables.[22] Similar point estimates and confidence bounds are provided when the welfare impacts are computed using the linear specification and when the nonlinear model is used, but a representative consumer approach is used to compute the welfare changes. These represent the second and third alternatives in Figure 2. The fourth and fifth alternative estimators correspond to the theoretical and computable bounds given by equations (10) and (15) above.

Several results emerge from Figure 2. First, in this application, the uncertainty regarding the GEV estimate of welfare is substantial, with the 95% confidence bound ranging from $9 to $25, encompassing both point estimates using the linear model and representative consumer approaches. Second, the representative consumer approach closely approximates the GEV sampler estimates. Both the point estimate of welfare ($16.51) and the confidence bounds are within six percent of the corresponding GEV sampler estimates. Finally, it is clear that the difficulties in computing the upper end of McFadden's theoretical bounds (specifically P_k^1 of equation 9) significantly reduces their information content. The theoretical bounds are relatively tight ($15 to $18), while the alternative computable bounds are considerably wider, only narrowing the compensating variation to lie in the range from $15 to $23. The appeal of these computable bounds is further reduced when it is recognized that the bounds themselves are uncertain, depending upon the estimated parameters of the model. Confidence intervals around the computable bounds are likely to blanket all of the previous alternatives.

Figure 3 provides a comparable set of result when the MNL error structure is used. The results parallel those in Figure 2. The point estimate of E(*cv*) for the representative consumer approach ($17.41) is virtually identical to the corresponding GEV sampler estimate, while the linear model's estimate is almost 17% larger. However, we again find that the difference in estimating the welfare gains by using the linear specification is swamped by the size of the confidence intervals surround each of the welfare estimates.

Finally, one limitation of our analysis thus far is that it has focussed attention on E(*cv*) averaged over the entire sample. There are two related problems here. First, to the extent that our sample is not representative of the population of interest, this estimate of E(*cv*) will be misleading. At a minimum, a weighting scheme would be required. It is not immediately clear how the alternative welfare measures would perform in this case. Second, policy analysts are often interested not only in the aggregate welfare impact of a program, but also in how it affects specific segments of the population; e.g., low income households. The comparisons in Figures 2 and 3 may mask potentially important differences in the welfare measures predicted for these sub-populations.

Figures 4 addresses these concerns by providing E(*cv*) by income quartiles using McFadden's (1995) GEV sampler, the representative consumer approximation, and the linear approximation. Here we again consider a doubling of the catch rate for all the modes using the Generalized Leontief model and the charter nesting error structure, paralleling the results in Figure 3. Three results emerge. First, as one might expect, $E(cv)$ does vary by income level, starting out low at roughly \$15 per choice occasion, rising at first as income increases, and then falling back below \$15.[23] Second, the representative consumer approximation to $E(cv)$ continues to track closely the GEV sample estimate, even when we focus on specific income levels. Third, the bias in restricting preferences to be linear in income does vary by income level. For the lowest income quartile, all three methods yield roughly the same welfare estimates. For the remaining quartiles, however, the linear approximation yields welfare estimates that are 17 to 24% smaller than the corresponding GEV sampler estimates.

VII. Conclusions

This study has investigated the importance of nonlinear income effects in random utility models, with particular attention to welfare measurement. In addition to specifying a nonlinear structure for the deterministic portion of consumer preferences, using Generalized Leontief and Translog models to provide flexible approximations to any nonlinear utility function, three distinct errors structures were considered. The resulting models were used to study mode choice among California anglers and to compare and contrast the available approaches for computing (or approximately) welfare changes when nonlinear income effects exist. These approaches include a re-sampling scheme based upon McFadden's (1995) GEV sampler, a linear model, a representative consumer approach, and the computation of bounds on the welfare changes of interest. The approaches trade-off computational ease for potential bias in the resulting welfare measures or uncertainty regarding their exact values.

Our analysis of Californian sportfishing represents, to our knowledge, the first application of McFadden's GEV sampler. Several key empirical results emerge. First, our findings (highlighted in Table 2) suggest that, in this application, there are more differences in the point estimates of welfare due to changes in assumed error distribution (e.g., multinomial logit versus nested logit) than due to the introduction of nonlinear income effects. Second, the consistent welfare estimates provided by the GEV sampler are not substantially different from the simpler linear and representative consumer approximations, particularly when the stochastic nature of the underlying parameter estimates is considered. Finally, while the computable bounds are both readily constructed and allow for nonlinear income effects, they do not provide tight bounds on the welfare estimates, even when one ignores the uncertainty of the underlying parameter estimates. Clearly, analysts must be cautious in drawing too strong of inferences from the results of this one data set. First, additional empirical examples are needed to determine the robustness of our findings. Alternatively, a Monte Carlo analysis, investigating those characteristics of consumer preferences that would widen the gap between the alternative welfare estimators, would be a natural direction for future research. However, we believe these results provide a useful point of departure. Second, while the differences among the welfare estimates with and without nonlinear income effects are generally small, they may represent a significant sum of money in actual policy settings, making the additional effort required to employ nonlinear specifications worthwhile in some circumstances.

In addition to providing an empirical comparison of alternative functional forms and error structures, we have also advanced the understanding of welfare measurement in discrete choice models by providing computable bounds based on McFadden's theoretical bounds. identifying cases in which those bounds are uninformative. and refining the procedures for computing the bounds themselves.

Table 1:
Parameter Estimates

Functional Form	Nesting Structure	β_{10}	β_{20}	β_{11}	β_{22}	β_{12}	θ	Log Likelihood
Linear	MNL	--	--	.02**	.95**	--	1.00	-1311.98
				(.00)	(.09)			
	Charter	--	--	.01**	.41**	--	.34[a]	-1235.17
				(.00)	(.10)		(.04)	
	Private	--	--	.01**	.85**	--	.62[a]	-1300.59
				(.00)	(.08)		(.06)	
Generalized Leontief	MNL	1.38**	1.99**	.01*	.47*	-.02*	1.00	-1303.91
		(0.45)	(.62)	(.00)	(.23)	(.01)		
	Charter	1.07**	-.00	.00	.70**	-.01	.31[a]	-1223.24
		(0.26)	(0.44)	(.00)	(.17)	(.01)	(.04)	
	Private	1.05*	.96*	.01*	.67**	-.01	.66[a]	-1295.28
		(0.42)	(.49)	(.00)	(.17)	(.01)	(.06)	
Translog	MNL	-40.90**	2.05**	6.71**	.14**	-.14**	1.00	-1297.47
		(2.52)	(.37)	(.39)	(.01)	(.044)		
	Charter	-77.70**	.44	7.37**	.05**	-.02	.34[a]	-1222.23
		(16.54)	(.25)	(1.24)	(.01)	(.03)	(.04)	
	Private	-33.24*	1.27**	5.05**	.11**	-.06	.64[a]	-1285.01
		(14.25)	(.29)	(.99)	(.01)	(.03)	(.06)	

**Statistically different from zero at a 1% significance level
*Statistically different from zero at a 5% significance level
[a]Statistically different from one at a 1% significance level

Table 2
Point Estimates of Welfare Impacts Using GEV Sampler

a. Doubling Prices

Nesting Structure	Linear	Generalized Leontief	Translog
MNL	-47.71	-47.53	-49.56
Charter	-48.79	-48.80	-48.90
Private	-48.42	-48.20	-49.74

b. Doubling Catch Rates

Nesting Structure	Linear	Generalized Leontief	Translog
MNL	20.33	17.41	15.89
Charter	14.15	16.95	7.95
Private	26.79	23.72	18.83

c. Loss of Shore Modes

Nesting Structure	Linear	Generalized Leontief	Translog
MNL	-35.89	-35.24	-35.27
Charter	-22.91	-21.79	-22.49
Private	-30.73	-30.84	-28.83

Figure 1
Alternative Nesting Structure

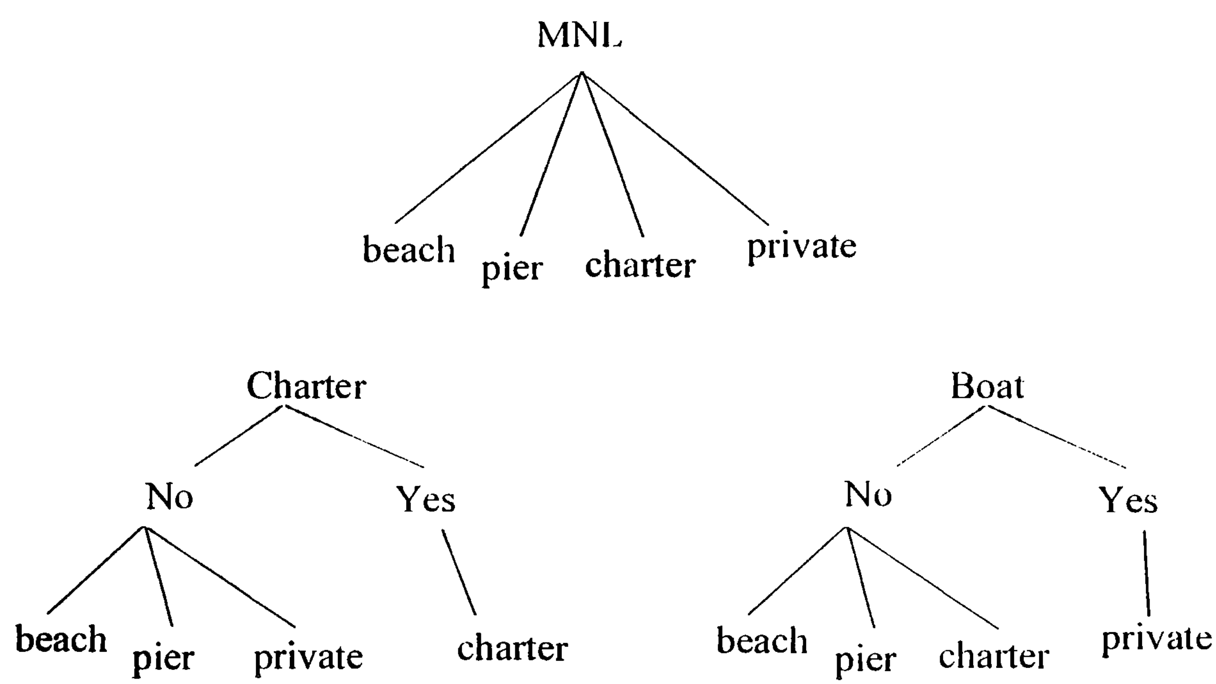


Figure 2
Alternative Welfare from Doubling Catch Rates - Charter Model

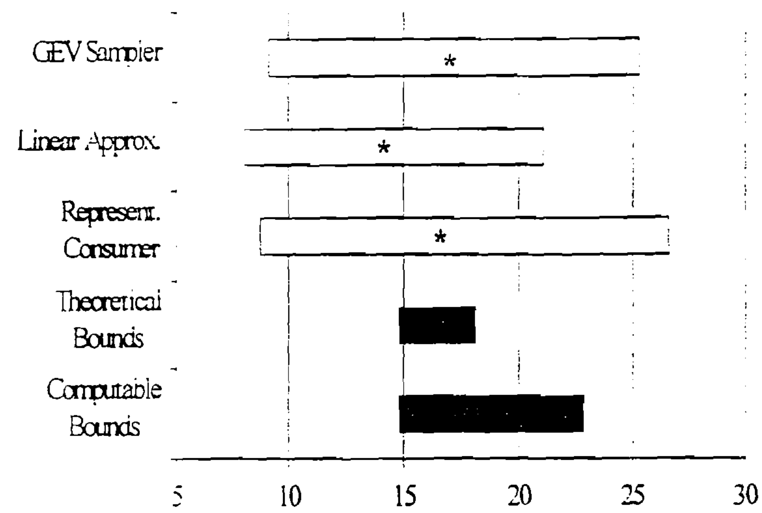


Figure 3
Alternative Welfare Impacts from Doubling Catch Rates - MNL Model

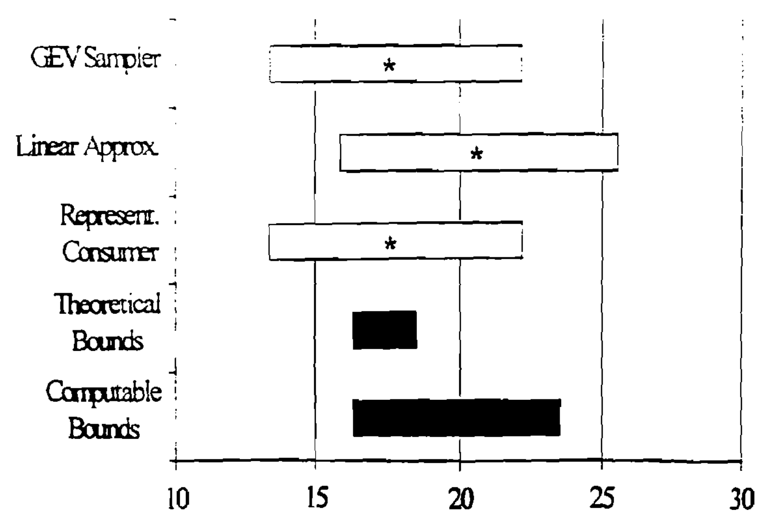
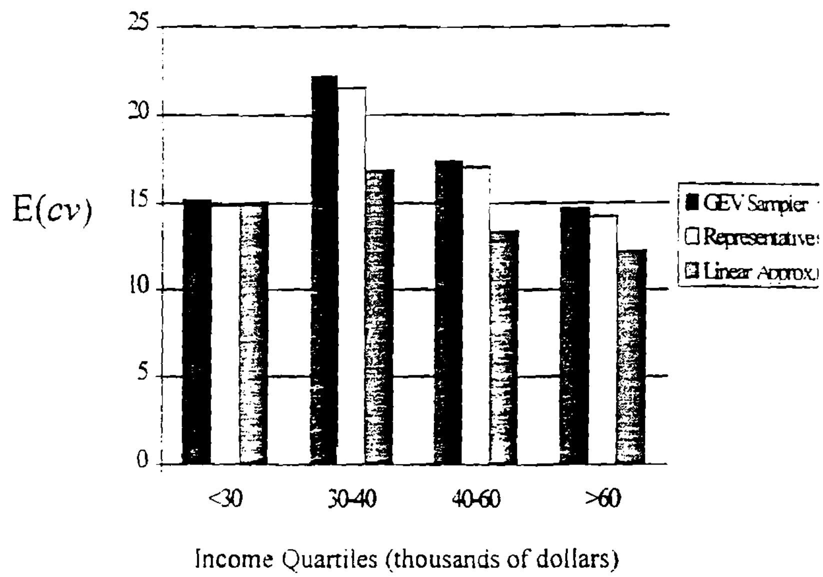


Figure 4
Alternative Welfare Predictions by Income Quartile - Charter Model



REFERENCES

Ben-Akiva, Moshe, and Steven R. Lerman. *Discrete Choice Analysis: Theory and Applications to Travel Demand* (Cambridge, MA: M.I.T. Press, 1985).

Bockstael, Nancy E., and Ivar E. Strand. "The Effect of Common Sources of Regression Error on Benefit Estimates." *Land Economics* 63 (Feb. 1987): 11-20.

Börsch-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 (New York: Springer-Verlag, 1987).

Börsch-Supan, Axel, and Vassilis A. Hajivassiliou. "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models." *Journal of Econometrics* 58 (1993):347-368.

Cameron, Trudy A., "A Nested Logit Model of Energy Conservation Activity by Owners of Existing Single Family Dwellings," *Review of Economics and Statistics* 67 (2, 1985):205-211.

Chipman, John S., and James C. Moore, "Compensating Variation, Consumer's Surplus, and Welfare," *American Economic Review* 70 (1980):933-949.

Chipman, John S., and James C. Moore, "Acceptable Indicators of Welfare Change," in John S. Chipman, Daniel McFadden, and Marcel K. Richter, eds., *Preferences, Uncertainty, and Optimality* (Boulder: Westview Press, 1990), pp. 68-120.

Daly, Andrew J. and Stan Zachary, "Improved Multiple Choice Models," in David Hensher and Quasim Dalvi, eds., *Identifying and Measuring the Determinants of Mode Choice* (London: Teakfield. 1979), pp. 335-57.

Diamond, Peter, and Daniel McFadden. "Some Uses of the Expenditure Function in Public Finance," *Journal of Public Economics*, (1974):3-21.

Domencich, Thomas, and Daniel McFadden. *Urban Travel Demand: A Behavioral Analysis* (Amsterdam: North Holland, 1975).

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

Gertler, Paul, and Paul Glewwe, "The Willingness to Pay for Education in Developing Countries: Evidence from Rural Peru," *Journal of Public Economics*, 42 (1990):251-275.

Geweke, John, "Bayesian Inference in Econometric Models Using Monte Carlo Integration," *Econometrica* 57 (1989):1317-1339.

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

Hanemann, W. Michael, "Welfare Evaluations with Discrete Models." Working Paper No. 792, Giannini Foundation of Agricultural Economics (May 1996).

Hausman, Jerry A., Gregory K. Leonard, and Daniel McFadden, "A Utility-Consistent Combined Discrete Choice and Count Data Model: Assessing Recreational Use Losses Due to Natural Resource Damage," *Journal of Public Economics* 56 (Winter, 1995):1-30.

Just, Richard. E., Darrell L. Hueth, and Andrew Schmitz. *Applied Welfare Economics and Public Policy* (Englewood Cliffs, N.J., Prentice Hall, 1982).

Kling, Catherine L., and Cynthia J. Thomson. "The Implications of Model Specification for Welfare Estimation in Nested Logit Models." *American Journal of Agricultural Economics* 78 (1996):103-114.

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior." In Paul Zarembka (ed.), *Frontiers Econometrics* (New York: Academic Press, 1973).

McFadden, Daniel, "The Measurement of Urban Travel Demand," *Journal of public Economics*, (1974):303-328.

McFadden, Daniel, "Modeling the Choice of Residential Location." In Anders. Karlqvist et al. (eds.). *Spatial Interaction Theory and Planning Models* (Amsterdam: North Holland, 1978) 75-96.

McFadden, Daniel "Econometric Models of Probabilistic Choice," in Charles F. Manski and Daniel L. McFadden (eds.), *Structural Analysis of Discrete Data* (Cambridge, MA: MIT Press, 1981).

McFadden, Daniel "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica* 57 (1989):995-1026.

McFadden, Daniel "Computing Willingness-to-Pay in Random Utility Models," mimeograph, Department of Economics, University of California, Berkeley, 1995.

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

Pollak, Robert, and Terence Wales, "The Likelihood Dominance Criterion," *Journal of Econometrics* 47 (1, 1991):227-242.

Shonkwiler, J. Scott, and W. Douglass Shaw, "Shaken, Not Stirred: A Finite Mixture Approach to Analyzing Income Effects in Random Utility Models," mimeograph, Department of Applied Economics and Statistics, University of Nevada, Reno, January 1997.

Shaw, W. Douglass, and Michael T. Ozog, "Modeling Overnight Recreation Trip Choice: Application of a Repeated Nested Logit Model," working paper, 1997.

Small, Kenneth A., and Harvey S. Rosen, "Applied Welfare Economics with Discrete Choice Models," *Econometrica* 49 (1981):105-130.

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

Train, Kenneth. E., Daniel L. McFadden, and Moshe Ben-Akiva, "The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices," *Rand Journal of Economics* 18 (1, 1987):109-123.

Yen, Steven, and Wiktor Adamowicz, "Participation, Trip Frequency and Site Choice: A Multinomial-Poisson Hurdle Model of Recreation Demand," *Canadian Journal of Agricultural Economics*, 42(1994):65-76.

VIII. Footnotes

[1]Important exceptions are the papers by Morey, Rowe and Watson (1993) and Shaw and Ozog (1997) on recreation demand and Gertler and Glewwe (1990) modeling the demand for schooling.

[2] Another consequence of assuming a constant marginal utility of income, as noted by McFadden (1995, p.10), is that it is then "...possible to aggregate preferences into a social preference that generates the market demand functions using Roy's Identity." See Chipman and Moore (1980,1990).

[3] Although we assume that the consumer is constrained to choose a single unit of the discrete good, the model can be generalized to allow multiple units.

[4] The indirect utility function is "conditional" on the choice of alternative j.

[5] As equation (3) implies, we are assuming that $\varepsilon^0 = \varepsilon^1$; i.e., that the preference heterogeneity is invariant with respect to the policy scenario. As McFadden (1995, p. 4) notes, without this assumption, the welfare impact of policy changes would no longer be well-defined and identifiable.

[6] Of course, depending on the application, the analyst may be interested in other moments of the distribution of compensating variations.

[7] Shonkwiler and Shaw (1997) have recently suggested an fourth alternative, using a finite mixture model to estimate consumer preferences that is piecewise linear in income.

[8] In our application below, numerical bi-section was used to solve for cv^i.

[9] The dissimilarity coefficient, denoted θ below, corresponds to the inverse of McFadden's (1995) "s".

[10] A procedure for constructing confidence bounds on the mean compensating variation estimates is outlined in the results section below.

[11] The exact form of the expected utility function depends upon the nesting structure assumed in the GEV distribution and are excluded here for the sake of space. See, for example, Morey (1994) for detailed expressions.

[12] Hanemann (1996) makes a similar observation.

[13] A second rationale for considering this representative consumer approach emerges if one considers the ε's as capturing individual uncertainty rather than heterogeneity of preferences across individuals. In this situation, equation (7) reflects the compensation calculation that would be undertaken by a risk neutral individual. This line of reasoning parallels the arguments put forth by Bockstael and Strand (1987) in the context of continuous demand systems.

[14] McFadden provides the intuition for this result by noting that if the consumer can move from the previously chosen alternative after a quality/price change, then the compensation needed to maintain the original level of utility may well be smaller than if forced to stay with the original choice. Likewise, the flexibility of having chosen another alternative prior to the price/quality change means that the consumer might need a smaller compensation than the one needed if they are forced to the finally chosen alternative.

[15] McFadden notes this point, but does not elaborate on it nor does he suggest a solution.

[16] Note that the computable lower bound and McFadden's theoretical lower bound are the same.

[17] We investigated a third GEV distribution that grouped the beach and pier alternatives together and the charter and private boat alternatives. However, the empirical results associated with the two structures reported in the paper dominated this structure based on both goodness-of-fit tests and consistency with utility maximization criteria.

[18] Note, however, that for any given error assumption, the Translog models provide slightly higher likelihood values than those from the Generalized Leontief.

[19] The choice of T = 1000 was selected on the basis of a Monte Carlo experiment in which the process of estimating $E(cv)$ using T iterations and the linear charter model was repeated 100 times. This exercise was conducted using various choices of T. The simulation results indicated that the estimated mean compensated variation changed little over the 100 trials once T exceeded 500, with the standard deviation of $E(cv)$ over the 100 trials reduced to less than $.05 by the time T = 1000. Since the charter model represents the extreme specification in terms of its departure from the MNL model, McFadden's (1995) simulation results would suggest that the GEV simulator would yield even more accurate welfare predictions for the private and MNL alternatives.

[20] The calculations reported in this paper were conducted using GAUSS, version 3.11, on a 200 MHz Pentium Pro IBM-compatible PC with 32M of RAM. While the calculation of *each* point estimate in Table 2 required only 15 minutes on this system, confidence bounds reported in Figures 2 and 3 below *each* required approximately 48 hours to construct.

[21] In constructing this figure, we assume that the correct model is the GL functional form with charter nested logit structure. The Generalized Leontief was chosen for further study

since it nests the linear model. An alternative to the charter nesting structure is considered in Figure 3.

[22] These confidence bounds were constructed by means of simulation, using the asymptotic distribution of the maximum likelihood parameter estimates to reflect the uncertainty in the model coefficients. 500 coefficient vectors were randomly drawn from the asymptotic distribution of the estimates of (β, θ). For each of these parameter draws, the GEV welfare estimate of $E(cv)$ was constructed. The 95% confidence bounds in Figure 2 reflect the middle 95% of the resulting estimates, dropping the smallest and largest 2.5% of the values.

[23] One might, at first, expect that the linear model would yield the same estimate for $E(cv)$ for each income quartile, given that this model assumes a constant marginal utility for both catch rates and income. However, as income changes, so do the travel costs associated with visiting a given site (since they depend in part on wage rates).

A Unified Theory of Recreation: Common Ground for the Random Utility and Hedonic Travel Cost Models Ψ

Linwood Pendleton
Assistant Professor of Economics
The University of Southern California
Los Angeles, CA

and

Robert Mendelsohn
Edwin Weyerhaeuser Davis Professor
Yale School of Forestry and Environmental Studies
360 Prospect Street
New Haven, CT 06511

Ψ This research was funded by the US Department of Energy with data collection support from the U.S. Forest Service. Special thanks to John Houghton and Tom Holmes for their support.

Abstract

This paper demonstrates that the random utility and hedonic methods emanate from the same utility theoretic foundation. Many of the differences in applications between the two techniques follow from *a priori* assumptions made by practitioners, not inherent differences in the theory of the methods. A theoretically consistent comparison of the two approaches is conducted valuing the quality of wilderness areas of the Southeastern United States. (*JEL* C25, Q23, Q26)

Finding Common Ground Among the Random Utility and Hedonic Travel Cost Models

Micro-economic theory began as an attempt to describe, predict and value the demand and supply of consumption goods. Quality was largely ignored in initial theoretical treatises as goods were assumed to be homogeneous. Over the last two decades, economists have started to apply the lessons from the theory of demand for goods to the demand for quality and more recently to the demand for recreational quality. Two distinct and seemingly inconsistent paths for incorporating quality into recreational site choice have emerged: the hedonic travel cost method (HTC) and the discrete choice random utility methods (RUM). The hedonic method views site attributes as though they were individual goods which just happen to be bundled together in a single purchase. The discrete choice approach models site choice based on a limited number of sites, all of which have different qualities.

Because the mathematical derivations for the hedonic [20] and random utility models (RUM) [14] are sufficiently distinct, many practitioners do not recognize that both models are based on a common utility theoretic foundation. In this paper, we develop a general theory of quality which we show is the basis for both models. Practitioners of the two methods, however, have often made different *a priori* assumptions about utility when applying the methods. For example, many studies using the RUM method have assumed linear utility functions while studies using the hedonic method frequently rely on quadratic utility functions. Further, the applications of both methods have made different assumptions about the error terms implicit in each method. In order to compare the two methods, it is important to make consistent assumptions about utility in both applications. Previous comparisons of the two models fail to ensure theoretic consistency between the two methods [4].

In the first section of this paper, a unified theory of the demand for quality is developed. The hedonic travel cost (HTC) and RUM methods are shown to be consistent with this unified theory. The paper uses two different functional forms for utility (linear and quadratic) in order to operationalize the theory and demonstrate the implications for both the hedonic and RUM approaches. The paper also offers a modification of the HTC that models site substitution following a change in site quality. An empirical example is then developed to compare the two methods and to demonstrate the implications of choosing linear versus quadratic functional forms for utility. The example measures the recreation value of forest attributes in the US Forest Service Wilderness Areas of the Southeastern United States. The empirical example indicates that the functional form of utility is important.

1. A Unified Theory of Quality

A change in the price of a good or a forced change in the consumption of a good causes a concomitant change in the well-being of the consumer. Utility theory measures the change in a consumer's well-being by calculating changes in the utility of the consumer and then converting these utility measures to some common numeraire. If the utility function is known, then the exact welfare measures of the change can be calculated directly. When the parameters of the utility function are unknown, the utility function and welfare measures can be derived from demand functions. If quality directly enters the consumer's utility function, the demand for quality is analogous to the demand for quantity. In this section, we demonstrate the link between the demand for quality and utility. We show that the hedonic and RUM models offer two distinct, but consistent ways to estimate these utility constructs.

1.1 The Hedonic Travel Cost Method

The theoretic links between the demand for goods and utility are now a well established part of basic micro-economic theory (see for example [23], [8], [13], and [10]). Without loss of generality, we extend this link to quality. We begin by considering a set of Hicksian demand functions for a vector of attributes (qualities), **Z**, described by a vector of attribute prices, **P**, utility u, and an estimation error term, ϕ.

$$(1) \qquad Z=h(\mathbf{P},u,\phi).$$

In the case of recreation demand, the price is not a market price, but an implicit price. This implicit price is found by estimating the hedonic price function. The hedonic price function is the empirical estimation of the hedonic price frontier in which sites that are visited are considered to be on the frontier and all other sites are considered to lie within the frontier. The cost of accessing any site on the frontier is a function of the attributes of that site. Formally, the hedonic price function is

$$(2) \qquad C(\text{site } j) = \text{fn}(\mathbf{Z}_j)$$

and the vector of implicit prices for the site attributes is given by the gradient of (2)

$$(3) \qquad \mathbf{P}=d\mathrm{C}/d\mathbf{Z}.$$

Given (1), we can find a set of inverse demand functions:

$$(4) \qquad \mathbf{P}=\mathrm{h}^{-1}(\mathbf{Z}, u, \phi).$$

Here **P** reflects the marginal value that the consumer would pay for an incremental unit of quality. We derive the consumer surplus associated with the consumption of **Z*** by taking a line integral of (4) from **Z**=0 to **Z**=**Z*** minus the costs of purchasing **Z***:

$$(5) \qquad \mathrm{CS} = \int_0^{\mathbf{Z}^*} \mathrm{h}^{-1}(\mathbf{Z},u)d\mathbf{Z} - \mathrm{C}(\mathbf{Z}^*)+\mathrm{g}(\phi)$$

Generally, practitioners take the expectation of $g(\phi)$ to be zero, but the exact structure of the error term in (5) depends on the nature of the error (e.g. omitted variables or measurement error) and whether the consumer surplus function is the direct integral of (4) (see [1]) or the recovered Hicksian consumer surplus (see [10]). Since $h(\mathbf{Z}, u, \phi)$ is a Hicksian demand, the consumer surplus measure in (5) is an exact measure of the welfare associated with **Z***; it is also a money metric utility function. Note that the definition of (5) allows for nonlinearity in the price schedule of **Z**.

One criticism of the hedonic method is that it estimates Marshallian demand, not the Hicksian demand function (1). Substituting a Marshallian demand function for (1) yields an inexact measure of consumer surplus in (5). However, Hausman [10] shows that an exact welfare measure can be recovered directly from the Marshallian demand function. Furthermore, when the assumed utility function is linear in income, the Marshallian demand is identical to the Hicksian compensated demand. Perhaps most importantly, the consumer surplus measures derived from Marshallian demand functions are approximately correct in almost all applications of this method (see [24]).

The simple demand relations (1) and (4) and the consumer surplus definition (5) are linked. Each of these relations defines the others. For instance, Englin and Mendelsohn [6], following LaFrance [13], show that the integration of the standard linear in attributes demand functions of the hedonic travel cost method yields a utility function that is quadratic in attributes and linear in income. Knowing any one of equations (1) through (5) gives all the information needed to calculate the welfare change that occurs when a consumer faces a change in **Z**. Without loss of generality, consider a type of recreation that can be described completely by a single attribute, z. If the single attribute, z, is available to consumers in fixed bundles at different recreational sites with each bundle i containing different qualities (i.e. quantities of z_i) available at cost $C(z_i)$, then equation (5) permits the calculation of welfare loss that is associated with a change in the quantities of z in one or more sites. Suppose a consumer chooses from a set of sites (where each site represents a specific bundle of quality). Now, suppose the amount of z available at each site changes from z^0 to z^1 across the n sites. The welfare value associated with this change is: [1]

$$(6) \qquad CS = \int_{z_i^0}^{z_j^1} h^{-1}(z,u)dz - C(z_j^1) + C(z_i^0) + g(\phi^1) - g(\phi^0)$$

where the subscript represents the site chosen, the superscript refers to the state of the world (i.e. before the change, 0, or after the change, 1), z_i^0 is the original level of consumption of attribute z at site i at cost $C(z_i^0)$ and z_j^1 is the new level of consumption of attribute z at site j at cost $C(z_j^1)$. If the consumer originally visits site *a* with attribute z_a^0 at cost $C(z_a^0)$ and continues to visit site *a* (i.e. i=j) at the same cost but now with z_a^1, then the exact welfare change is:

$$(7) \quad \Delta CS = \int_{z_a^0}^{z_a^1} h^{-1}(z,u)dz .$$

If the consumer no longer visits site *a*, then we must first determine which new site the consumer would choose. Traditional applications of the hedonic method failed to predict how consumers would choose amongst recreational sites following a quality change ([6]). A rational consumer ought to choose the site that maximizes utility. From the analyst's perspective, the consumer's problem is a probabilistic one in which the probability of choosing a site becomes

$$\text{(9)} \qquad \text{Prob(choose site j)} = \text{Prob}[CS(j)] \geq \text{Prob}[CS(\backslash j)].$$

If we know the distribution for $g(\phi)$, then we can calculate the probability that any individual will choose site j. Including this extra step extends the hedonic travel cost analysis to the prediction of site substitution. We call this the substitution-adjusted hedonic method (SAH). We develop the substitution adjusted method in a later section.

1.2 The Random Utility Method

The random utility method begins where the hedonic method ends. The method focuses upon choosing from a set of discrete sites each of which embodies a vector of attributes (qualities). Following McFadden [14], the problem is equivalent to (9) but now the consumer chooses a site to maximize their conditional utility

$$\text{(10)} \qquad \text{Prob(choose site j)} = \text{Prob}[U(\mathbf{Z}_j, X) + \varepsilon_j] \geq \text{Prob}[U(\mathbf{Z}_{\backslash j}, X) + \varepsilon_{\backslash j}]$$

such that. $Y= [HX+C(\mathbf{Z}_j)]$.

The conditional utility of the RUM may have the same functional form as the CS from the hedonic approach provided CS derived from the Hicksian demand. The random utility function consists of a deterministic core, $U(\mathbf{Z}_j,X)$, and a random component, ε_j. This random utility is a function of the attributes, $\mathbf{Z}_j$, of the site chosen, j, and all the remaining goods, X, that can be consumed. H is the price of other goods X, ε is a random variable and Y is income. If we assume, without loss of generality, that the price, H, of other goods X is 1, then we can substitute $HX= Y-C(\mathbf{Z}_j)$ into the utility function in equation (12) in which case the consumer chooses a site to maximize their random utility:

(11) $U(\mathbf{Z}_j, Y-C(\mathbf{Z}_j)) +\varepsilon_j$

Note that (11) is now a conditional random utility, conditional upon choosing to visit a site. As we showed with the hedonic method, any functional form for the deterministic portion of the random utility, U, implies a demand function for Z.

Our ability to calculate the welfare change that would result from a change in quality depends on our ability to estimate the parameters of at least one of the equations given in (1) through (5). The hedonic method estimates the parameters of the demand functions (1) and the RUM estimates the parameters of the utility function (5). Both rely upon the information inherent in the consumers' choice of quality[2].

Note that since (5) is the integral of (4), for the two approaches to be theoretically consistent, the functional form for (5) must be the integral of (2). In Section 2, we illustrate this point with two examples, a linear and a quadratic utility function.

1.3 Calculating Expected Welfare

Ideally, we want to calculate a measure of welfare change that demonstrates the compensation or payment required to maintain utility with and without an environmental change. Measures of equivalent or compensating variation are well developed in the literature on welfare analysis. With the RUM, we do not know the exact utility of our representative consumer. Instead, we estimate a random utility. This means that we calculate expected utility and thus our estimate of welfare change also is an expected measure. Formally, the expected utility of a representative consumer is

$$(12)\quad E[\mathrm{U}] = \sum_{j=1}^{n} \mathrm{Prob(site\ j)} * \mathrm{U}_j$$

where the probability of choosing any one of the n sites is given by

$$(13)\quad \mathrm{Prob(site\ j)} = \int_{-\infty}^{\infty} \prod_{j \neq i} \mathrm{F}(\varepsilon_j + \mathrm{U}_j - \mathrm{U}_i) f(\varepsilon_j) d\varepsilon_j$$

where F(•) and f(•) are cumulative and probability distribution functions respectively. A change in quality at one or more sites causes a change in expected utility not only because the deterministic portion of utility, U, changes at the affected site(s) but also because the probability of choosing each site changes. Traditionally, RUM practitioners assume a generalized extreme value distribution for ε and thus the change in expected utility can be found by

$$(14) \quad E[\Delta U] = \left\{ \ln\left[\sum_j \exp\{U(\mathbf{Z}_j^1)\} \right] - \ln\left[\sum_j \exp\{U(\mathbf{Z}_j^0)\} \right] \right\},$$

where the superscripts represent states of the world in terms of site quality. When the marginal utility of income is assumed to be constant (usually the case in RUM applications) the change in expected utility is converted to a money metric numeraire by dividing through by the marginal utility of income, λ, giving an expression for the expected change in welfare [9]. In this case, since the marginal utility of income is constant, the welfare measure is equal to compensating variation which is equal to equivalent variation which is equal to consumer surplus. Formally the expression for change in welfare is

$$(15) \quad E[\Delta CV] = \frac{1}{\lambda} \left\{ \ln\left[\sum_j \exp\{U(\mathbf{Z}_j^1)\} \right] - \ln\left[\sum_j \exp\{U(\mathbf{Z}_j^0)\} \right] \right\}.$$

Earlier in the paper, we showed that a money metric utility function also can be derived from the hedonic demand functions. Like the random utility models, the utility functions derived from the estimated hedonic demand functions also contain a random term. From (5) we have

$$(16)\ \mathrm{CS} = \int_{0}^{\mathbf{Z}^*} \mathrm{h}^{-1}(\mathbf{Z},u)d\mathbf{Z} - \mathrm{C}(\mathbf{Z}^*) + \mathrm{g}(\phi).$$

Remember that when $h^{-1}(\mathbf{Z}, u)$ is the inverse Hicksian demand function, CS is a money metric utility function. Also, when there are no income affects (i.e. the marginal utility of income is constant), the Marshallian consumer surplus derived in (16) is identical to the Hicksian measure and also is a money metric utility function. Following the derivation of welfare change for the RUM, we could also develop an expected welfare measure using the hedonic results. Specifically, the expected money metric utility (consumer surplus) for a representative consumer would be

$$(17) \qquad \mathrm{E}[\mathrm{C}S] = \sum_{j=1}^{k} \mathrm{Prob(site\ j)} * \mathrm{CS}(\mathbf{Z}_j).$$

Also following from above, the probability of visiting any site is

$$(18) \qquad \mathrm{Prob(site\ j)} = \int_{-\infty}^{\infty} \prod_{j \neq i} \mathrm{F}[\mathrm{g}(\phi_j) + \mathrm{CS}(\mathbf{Z}_j) - \mathrm{CS}(\mathbf{Z}_i)] * f(\mathrm{g}(\phi_j))d\mathrm{g}(\phi_j)$$

Obviously, integrating over the distribution of $g(\phi_j)$ could be quite difficult and a closed form solution may not exist for certain distributions of $g(\phi_j)$. Nevertheless, as a first approach, we investigate the impact on welfare calculations of an expected hedonic welfare measure by assuming arbitrarily that $g(\phi_j)$ has a generalized extreme value (logistic) distribution. In this case, the expected welfare measure for a change in quality is

$$(19)\quad E[\Delta CS] = \left\{ \ln\left[\sum_j \exp\{CS(\mathbf{Z}_j^1)\} \right] - \ln\left[\sum_j \exp\{CS(\mathbf{Z}_j^0)\} \right] \right\}.$$

2. **Implications of Utility Functional Form**

2.1 Linear Utility

Utility functions that are linear in both attributes and income (cost) are used commonly in applications of the RUM to recreational quality (e.g. [2], [17], [18], and [12], and [11]). The standard deterministic core of the linear utility function is

$$(20)\quad U_j = \beta \mathbf{Z}_j + X \text{ subject to } Y = HX + C(\mathbf{Z}_j),$$

where subscript j refers to site j, $\mathbf{Z}_j$, is the vector of quality attributes that describe site j, $C(\mathbf{Z}_j)$ is the cost of accessing site j with attributes $\mathbf{Z}_j$, and Y, H, and X are as before. If we assume that utility (20) is linear in income (all other goods), and that H is fixed and can be set arbitrarily to unity, then we can use the income constraint to substitute $Y\text{-}C(\mathbf{Z}_j)$ for X giving us:

(21) $U_j = \beta\mathbf{Z}_j + \lambda\ [Y\text{-}C(\mathbf{Z}_j)],$

where λ can be interpreted as the (constant) marginal utility of income. Equation (21) forms the deterministic core of the RUM in which the conditional random utility derived from choosing site j is

(22a) $v_j = U_j + \varepsilon_j$

(22b) $v_j = \beta\mathbf{Z}_j + \lambda[Y\text{-}C(\mathbf{Z}_j)] + \varepsilon_j,$

where ε_j is a random term. Most frequently, the RUM is estimated assuming a logistic or extreme value distribution for ε_j. The estimation of the RUM proceeds by a differences in utility specification in which the differences in utilities between sites also has the same distribution as the random term. In the differences in utilities approach, λY disappears from the utility function because income does not vary across sites and the marginal income of utility is assumed to be constant. In the econometric application of the RUM, the conditional random utility function becomes

(23) $v_j = \beta\mathbf{Z}_j - \lambda C(\mathbf{Z}_j) + \varepsilon_j.$

Note that (23) also is a conditional indirect utility function in price ($C(\mathbf{Z})$) and quality, $\mathbf{Z}_j$.

The deterministic portion of the linear utility function is not strictly "well-behaved" in the sense that it is not strictly concave. The linearity of the utility function means that the marginal value of any attribute remains the same for all levels of quality (i.e. the marginal value is constant).

$$(24) \qquad \frac{\partial U}{\partial \mathbf{Z}} = \beta,$$

where $\frac{\partial U}{\partial \mathbf{Z}}$ is a column vector of marginal utilities and β is a column vector of coefficients. If a single linear utility function is thought to apply to all consumers, then we assume that all consumers place the same marginal values on attributes, $\mathbf{Z}$, regardless of how much is purchased.

The linear utility function can be estimated by the hedonic method by estimating a single linear hedonic price function for all markets (origins).

A linear utility function may approximate a well-behaved non-linear utility function (i.e. where utility is strictly concave in $\mathbf{Z}$ and monotonically increasing) when consumption levels of $\mathbf{Z}$ vary over a narrow range. Clearly, if the underlying true utility function is nonlinear and a wide range of qualities are being considered, this approximation becomes more troublesome.

2.2 Quadratic Utility

Many applications of the hedonic method to recreational quality implicitly assume a utility function that is quadratic in its non-income arguments ([15], [21], [22]). More sophisticated applications of the HTC assume quadratic utilities that also contain cross-price terms (e.g. [6], [3]). The functional form for the deterministic core of the quadratic utility function is:

$$(25) \quad U_j = \tfrac{1}{2}(\mathbf{Z}_j - \alpha)'\beta^{-1}(\mathbf{Z}_j - \alpha) + \lambda X \text{, subject to } Y = HX + C(\mathbf{Z}_j),$$

where $\mathbf{Z}$ is a vector of site attributes, α is a vector of constants, and β is a matrix to be estimated. A well-behaved quadratic utility function requires that all elements of the vector α are positive and that the matrix β is negative semi-definite. The cross-price terms allow attributes to act as substitutes or complements.

With the quadratic utility function, it is theoretically possible to have oversatiation if a consumer faces a cheap (nearby) and over-abundant supply of a specific attribute. For some economists, the potential for negative prices (decreasing utility with increasing attributes) is sufficient reason to reject a quadratic utility functional form [7]. There are, however, two cases in which a quadratic utility function might be appropriate for the analysis of recreational quality. The first case is where the feasible consumption set is one in which all or most consumers have a utility that lies within the increasing range of the utility function. The second case is when consumers do not enjoy free disposal [5] and may be forced to consume a level of attributes that exceeds the level of complete satiation. For example a skier may happen to live near a ski area which has exceedingly large amounts of a normally desirable attribute such as deep snow. The consumer cannot sell off this overabundance and may be observed to occasionally travel further (pay more) to go to a site with less snow. The negative prices often found in applications of the HTC can reflect oversatiation. Results using the same data that follow in Section 3 and published in another paper [19] show that for hiking in the Southeastern United States, negative implicit prices are associated with attribute levels that are significantly higher than attribute levels where prices are positive.

The hedonic method estimates the parameters of the quadratic utility function by first estimating a hedonic price function for each origin in which $C(\mathbf{Z})$ is regressed upon $\mathbf{Z}$. Any functional form can be used in the regression. Using these hedonic prices, a system of seemingly unrelated demand functions is estimated

$$(26) \quad \mathbf{Z}=\alpha+\beta\mathbf{C}_Z+\phi.$$

where $\mathbf{Z}$, α, $\mathbf{C}_Z$ are the same vectors as before, ϕ is a vector of error terms, and β is a matrix. In order to integrate (26) back to (25) and to ensure that integration is path independent, it is necessary to constrain the cross-diagonal elements of β to be symmetric (the Slutsky conditions).

The RUM analysis can estimate the coefficients of the quadratic random utility function after expanding the vector notation of (25). A simplified form of the expanded utility would follow:

(27)

$$U = \left[\beta_1^{rum} z_1 + \frac{1}{2}\beta_2^{rum} z_1^2 + \ldots + \beta_n^{rum} z_n + \frac{1}{2}\beta_{n-1}^{rum} z_n^2 + \beta_{n+1}^{rum} z_1 z_2 \right] + \alpha^{rum} \ldots + \lambda\big(Y - C(\mathbf{z})\big) + \varepsilon$$

where the coefficients, α^{rum} and β^{rum} represent collected terms (i.e. the complex coefficients that result from the matrix multiplication in (25)). The income constraint is substituted in for all other goods X in (25). Unlike the hedonic estimation, there is no need to restrict cross-price terms since only one coefficient is estimated for each cross-attribute pairing. The constant, α^{rum}, cannot be estimated using the RUM and is irrelevant for welfare and utility calculations. As with the linear utility function (21), the income term, Y, is dropped in the standard RUM estimation.

3. **An Empirical Comparison**

Past comparisons between hedonic and RUM methods have made no attempt to make consistent assumptions about the underlying form of the utility function (e.g. [2] and [4]). In this section we estimate both linear and quadratic utility functions for the HTC and RUM methods.

3.1 Data

Data were collected on 4778 visitors to 46 trails in 20 different forest areas near the Smoky Mountains (see [19]). Visitor data came from permits collected by the USDA Forest Service (USFS) and an independent survey. We limit the data set to visitors from within 300 miles of the North Carolina and Tennessee border in order to focus the analysis on single purpose trips. The data were collected between 1992 and 1994. Trails were surveyed in non-wilderness areas, the State Park system, and the Great Smoky Mountain National Park.

Important trail attributes were identified by interviewing hikers and reading popular trail guides. Standard ecological techniques were used to measure these attributes along each of the 46 trails in the study. The set of trail attributes includes "basal area" (a measure of the size of trees and tree density), "elevation" (the maximum elevation of each trail), "riparian" (percent of trail along a creek), and "isolation" (measured as miles from the paved road to the trail head). Appendix A gives summary statistics for the trail attributes. In addition, the distance from each origin to a trailhead was calculated using the program ZIPFIP (USDA 1993). All distances are in one way miles.

3.2 The Methods

Both the RUM and HTC methods are estimated according to standard practice. We give a brief review of the estimation methods here.

The Hedonic Cost Function

We estimate the implicit price of trail attributes by regressing the total travel costs to sites visited, C(**Z**), on levels of environmental attributes at these sites. Because the geographic configuration of sites differs for every origin, a different hedonic price function is estimated for each origin. Using OLS, we estimate the hedonic price function only for those sites actually visited by residents of a given origin. It is assumed that sites that are not visited are not on the hedonic price frontier (i.e. these sites are inferior). We assume that the hedonic price function is linear:

$$(28) \quad C(\mathbf{Z}) = c_0 + C_1(\text{basal}) + C_2(\text{elevation}) + C_3(\text{riparian area}) + C_4(\text{isolation}) + \psi$$

where **Z** is a vector of quantities for the selected attributes (basal, elevation, riparian, isolation) and ψ is the estimation error. The coefficients, C_i, represent the implicit prices for the attributes. Because we run a different regression for each origin, a different vector of implicit prices, $\mathbf{C_z}$, , exists for each origin.

The coefficients of the hedonic cost function represent the implicit prices of attributes. These implicit prices represent the marginal value of any attribute. The linear in attributes utility function implies a constant marginal value for each attribute, regardless of the level of attributes consumed. Therefore, a single hedonic cost function also was estimated for all origins simultaneously. The coefficients of this "universal" linear hedonic cost function are consistent with the marginal values that would be derived from the linear utility function.

The Demand for Site Attributes

The second step in the hedonic travel cost analysis is to estimate the demand for site attributes based on the implicit prices faced by each visitor and the level of attributes chosen by each visitor. In this study, we estimate a system of demand functions that are linear in site attributes and socio-economic shift variables. Using data on all visitors, we estimate the following system of demand functions:

$$(29) \qquad \mathbf{Z}=\alpha+\beta\mathbf{C}_Z+\delta S \qquad +\phi$$

where **Z** is a vector of quantities for the selected attributes (basal area, elevation, riparian, isolation), $\mathbf{C}_Z$ is a vector of hedonic prices from the first stage regressions, S is a vector of socio-economic variables, ϕ is a vector of estimation errors and α, β and δ are respectively a vector and two matrices of coefficients to be estimated. The socio-economic shift variables are characteristics of each origin and are derived from U.S. 1990 census data (Hellerstein et al. 1993). Interestingly, we could not identify any socio-economic variables that significantly affected the demand for site attributes and so S was dropped from (29). Because the coefficient on income (an element of S) was not significantly different from zero, we conclude that the income elasticity of demand for forest attributes is zero and thus compensating variation, equivalent variation, and consumer surplus are equivalent.

The prices from the first stage and the quantities of site attributes chosen by hikers allows us to estimate the demand functions of equation (29). Because hikers from different origins face different prices, we treat each origin as a separate market. The existence of multiple markets allows the estimation to be specified and avoids the pitfalls common to single market hedonic applications (see [16]). We estimate equation (29) using a generalized least squares, seemingly unrelated regression procedure. We constrain the cross-prices of β to be symmetric in order to ensure that welfare measures are path independent.

The Random Utility Models

We estimate the RUM models using standard non-nested multi-nomial logit methods. All trails are included in the choice sets of individuals. We estimate a linear in attributes conditional random utility function

$$(30) \quad v_j=\beta^{\text{linear rum}}\mathbf{Z}_j + \lambda[Y-C(\mathbf{Z}_j)] + \varepsilon_j,$$

where $\mathbf{Z}_j$ is defined as before (i.e. $\mathbf{Z}_j=$ {basal area, elevation, riparian, isolation}). We also estimate a quadratic in attributes random utility function

$$(31) \quad v_j=\beta_1(\text{basal area})+ \beta_2(\text{basal area})^2+\beta_3(\text{elevation}) +\beta_4(\text{elevation})^2+\beta_5(\text{riparian})+ \beta_6(\text{riparian})^2+\beta_7(\text{isolation})+ \beta_8(\text{isolation})^2 + \varepsilon_j$$

where all of the coefficients, of course, refer only to the quadratic specification.

3.3 Econometric Results

We estimate two models: the hedonic travel cost (HTC) and the random utility model (RUM). We also explore two different functional forms for the utility function: linear and quadratic. The RUM estimation used a non-nested multinomial logit format. Details of the hedonic estimation are available in [19]. The estimated coefficients for the models with linear utility are presented in Table Ia and the coefficients for quadratic utility are presented in Table IIa.

The linear utility parameters for the HTC model suggest that basal area and elevation are both goods whereas isolation is an economic bad and creek is not relevant. The results from the linear RUM analysis suggest that both elevation and creek are bads whereas basal area and isolation are good. Although the basal area and isolation results are consistent with prior expectations, the remaining results from the RUM analysis seem inconsistent with the description of trail attributes in hiking books.

We can compare the degree to which each method estimates the same coefficients by examining the parameters of any one of the utility/demand relationships in (1) through (5). We compare the models by examining the parameters of the derived inverse demand functions because the errors of the parameters can be calculated easily. A Wald test shows that the linear HTC and RUM analyses do not estimate identical parameters for the inverse demand function (Table Ib). The marginal prices of both isolation and basal area in the HTC are significantly different from the marginal prices estimated by the RUM analysis.

In general, the results of the quadratic utility function are superior to the linear utility model. More coefficients are significant and have the expected sign and the models explain a greater fraction of the observed behavior. The quadratic utility parameters for the HTC model imply negative own price elasticities (downward sloping demand functions) for all four attributes. The cross price elasticities between basal area and both creek and isolation are positive implying these attributes are substitutes. Elevation also has a positive cross price elasticity with respect to isolation. The quadratic utility parameters for the RUM model yield similar results. The linear and quadratic coefficients for both basal area and creek have the expected sign. All interaction terms between attributes suggest that the attributes are substitutes. Neither model, however, performs exactly as expected. If the estimated coefficients are taken at face value, the HTC model implies that creek is a bad not a good. However, it should be noted that the coefficient on the interaction term between creek and elevation is not significantly different from zero. If we set this term equal to zero, then the hedonic method depicts creek as a normal good. In contrast, the RUM model suggests that the more isolation and the more elevation, the better the site becomes at an increasing rate.

Seventeen of the 20 common coefficients between the HTC and RUM models are of the same sign. Using a Wald test to compare the coefficients of the inverse demand functions (Table IIb) shows that 12 of the 20 coefficients are not significantly different at the 95% level. All of the coefficients, however, differ by at least one order of magnitude. All but one of the coefficients which were significantly different between the RUM and HTC models involved creek.

3.4 Welfare Estimates

For perspective, welfare estimates are given for changes in the levels of attributes at all trails and changes in the levels of attributes at a single trail (the Pleasant Garden Overlook in the Unaka Wilderness Area of Tennessee). Attribute changes are calculated for a change in the level of each attribute equal to 10% of the mean across all sites. All of the results are given as mean welfare changes, in dollars assuming that it costs $0.25 per mile traveled. The welfare results are proportional to the assumed travel cost so that the reader can easily adjust these figures for different travel cost per mile estimates.

We make two different welfare measurements using the HTC empirical estimates. In the analysis labeled HTC, we assume that people stay at the same site before and after the quality change. The welfare measure is (7). In the analysis labeled SAH the welfare is the expected welfare change given in (17). With the RUM analysis, an expected welfare measure also is used (15).

The calculation of CV under the SAH, and analogously the RUM, departs from the traditional HTC in that these methods allow the visitor to change sites. Because a visitor will only choose a new site if it makes him better off, relaxing the restrictive assumption of holding destinations constant should increase utility. Thus, we might expect that the SAH will yield larger (smaller) estimates for beneficial (detrimental) changes than HTC.

The welfare estimates for a 10% decrease in each attribute for all trails is presented in Table III. The results from the RUM linear utility model do not appear consistent with intuition. The RUM model predicts that decreases in elevation and creeks would improve the value of a trail. The HTC model also predicts one strange result; decreases in access to creeks would improve trip value. The quadratic results for the RUM are more unexpected than in the case of linear utility. Decreases in three attributes considered goods (basal area, elevation, and creek) are predicted to increase site value. The anomalous welfare results from the quadratic RUM are the result of the interaction terms, the coefficients of which are all negative. The HTC model, in contrast, predicts that a decrease in basal area, elevation, creek, and isolation would all reduce site value. The SAH model produces quite similar results. The SAH model predicts slightly smaller basal area effects and slightly larger elevation values but is otherwise the same as the HTC results.

In Table IV, we examine the welfare impact of a quality change at only a single site. A decrease in the level of attributes at one site should have only a marginal impact on visitor welfare because visitors can readily move to substitute sites. Methods which allow this substitution (e.g. the RUM and SAH) should predict smaller welfare effects from quality reductions at a single site. It is therefore reassuring that the results in Table IV for the SAH and the RUM models are substantially smaller than the results reported in Table III where every site underwent a change. For the quadratic RUM, the greatest welfare loss (-$0.14) occurs for a 10% loss in the level of basal area at the Pleasant Garden Overlook Trail. For the quadratic SAH, the maximum welfare loss (-$0.42) occurs for a 10% reduction in any attribute. The SAH values a 10% loss of any attribute the same since changes in all four attributes would result in the consumer shifting to another site. The results for the quadratic HTC are much larger. The maximum welfare loss for the HTC (-$3.70) for basal area is considerably larger than the results using the other methods. The assumption that visitors would stay at a site when its quality declines leads to a much larger estimate of welfare damages.

4. **Conclusion**

Unlike other studies that compare the hedonic and RUM methods (e.g. [4], [2]), this study compares the models under identical utility theoretic assumptions. This study shows that neither the hedonic nor the RUM models can be selected *a priori* simply because they are based on a better theoretical foundation. The hedonic and RUM models are consistent theoretically; each method simply estimates a different function in the chain of functions that links demand with utility and each method makes different assumptions about the error term.

The empirical comparison of the HTC and RUM indicates that practitioners need to take great care when making theoretical assumptions prior to estimation. Despite having consistent utility functional forms, the HTC and RUM produce strikingly different welfare results. The study also indicates that both the hedonic and

RUM models are highly sensitive to the functional form chosen for utility. The paper shows that linear in attributes utility functions imply restrictive assumptions about the way in which consumers value recreation quality. Analysts should be cautious when assuming that utility is linear.

The hedonic and RUM methods are both important tools in the valuation of quality change. Each method makes different demands on data and attributes to be estimated. The hedonic methods are well-suited to choices where attribute levels vary smoothly and continuously, and when there are abundant alternatives. The RUM methods lend themselves to choice sets which are limited with discrete attribute levels. The econometric estimation of both methods often require strong assumptions about supply and demand. Problems with the identification of supply and demand functions in the hedonic methods are well known. Less attention has been given to the restrictions implicit in the RUM. Nevertheless, both tools need to be used to learn more about the role of quality in consumer choice.

References

Bockstael, N. and I. Strand. 1987. The effect of common sources of regression error on benefit estimates. Land Economics. 63(1):11-20.

Bockstael, N., Hanemann, W.M. and C. Kling. 1987. Estimating the value of water quality improvements in a recreational demand framework. Water Resources Research. 23:951-960.

Brown, G. and R. Mendelsohn. 1984. The hedonic travel cost method. Review of Economics and Statistics. 66: 427-433.

Cropper, M., Deck, L., Kishor,N., and K. McConnell. 1993. Valuing product attributes using single market data: a comparison of hedonic and discrete choice approaches. Review of Economics and Statistics; 75(2):225-32.

Englin, J. 1986. Backcountry hiking and optimal forest management. p. 57. University of Washington, Seattle.

Englin, J and R. Mendelsohn. 1991. A hedonic travel cost analysis for valuation of multiple components of site quality: the recreation value of forest management. Journal of Environmental Economics and Management. 21:275-290.

Freeman, M. 1993. The Measurement of Environmental and Resource Values: Theory and Methods. Resources for the Future.

Green, J., Mas-Colell, and M. Whinston. 1995. Microeconomic Theory. Oxford University Press

Hanemann, M. 1985. Applied welfare analysis with discrete choice models. Working paper. University of California, Berkeley.

Hausman, J. 1981. Exact consumer's surplus and deadweight loss. The American Economic Review. 71(4):pp 662-676.

Hausman, J., Leonard, G. and D. McFadden 1995. A utility-consistent, combined discrete choice and count data model: assessing recreational use losses due to natural resource damage. Journal of Public Economics. 56: pp 1-30.

Kaoru, Y, V. Smith, and J Liu. 1995. Using random utility models to estimate the recreational value of estuarine resources. American Journal of Agricultural Economics. 77:141-151

LaFrance, J.T. 1985. Linear demand functions in theory and practice. Journal of Economic Theory. 37:137-166.

McFadden, D. 1978. "Modeling the choice of residential location" in A. Karlquist et al. (eds.) Spatial Interaction Theory and Residential Location, pp. 75-96.

Mendelsohn, R. 1984. An application of the hedonic travel cost framework for recreation modeling to the value of deer. Advances in Applied Microeconomics 3:89-101

Mendelsohn, R. 1987. A review of identification of hedonic supply and demand functions. Growth and Change. Winter. pp 82-92.

Parsons, G. and M. Kealy. 1992. Randomly drawn opportunity sets in a random utility model of lake recreation. Land Economics. 68(1):93-106

Parsons, G. and M. Needleman. 1992. Site aggregation in a random utility model of recreation. Land Economics. 68(4):418-33

Pendleton, L., Sohngen, B., Mendelsohn, R. and T. Holmes. In review. Measuring the Economic Value of Forest Attributes in the Southern Appalachian Mountains.

Rosen, S. 1974. Hedonic prices and implicit markets: product differentiation in pure competition. Journal of Political Economy. 82:34-55

Smith, V.K. and Y. Kaoru. 1987. The hedonic travel cost method: a view from the trenches. Land Economics. 63:779-792.

Smith, V. , Palmquist, R. and P. Jakus. 1991. Combining Farrell frontier and hedonic travel cost models for valuing estuarine quality. Review of Economics and Statistics. 73(4): 694-699

Varian, H. 1992. Microeconomic Analysis. W.W. Norton & Company. New York.

Willig, R. 1976. Consumer's surplus without apology. The American Economic Review. 66(4):pp 589-597.

Tables

Table Ia: The Estimated Parameters of the HTC and RUM: Linear Utility

HTC Results

		Constant	basal area	elevation.	riparian	isolation
	$C(\mathbf{z})$ =	66.2	0.199	5.81×10^{-3}	-1.96	-2.12
	(t-statistics)	(3.90)	(1.57)	(3.27)	(-0.216)	(-6.04)
	observations = 4778		corrected r^2=.0201			

RUM Results

		basal area	elevation	riparian		isolation	travel cost
	$v(\mathbf{z},C)$ =	2.57×10^{-2}	-4.99×10^{-5}	-0.513		0.103	-2.97×10^{-2}
	(t-statistics)	(21.6)	(-25.6)	(-6.64)		(24.0)	(-46.0)
	observations = 4778		percent sites correctly predicted		31.65		log likelihood= -13197

Table Ib: The Parameters of the Inverse Demand Functions: Uniform Linear Utility
(bold face indicates coefficients are not different at the 5% significance level)

	$C_{\text{basal area}}$	$C_{\text{elevation}}$	C_{riparian}	$C_{\text{isolation}}$
HTC	0.199	**5.81×10^{-3}**	**-1.96**	-2.12
RUM	0.866	**-1.68×10^{-2}**	**-17.3**	3.45
Wald Test	24.3	**0.335**	**2.86**	239

Table IIa: The Estimated Parameters of the HTC and RUM: Quadratic Utility
(t-statistics in parentheses)

HTC		basal area	elevation	riparian	isolation	
	constant	79.2	2990	0.284	5.73	
		(215)	(143)	(65.6)	(100)	
	$C_{basal\ area}$	-7.18	22.8	0.502×10^{-1}	0.652	
		(-11.8)	(0.689)	(8.95)	(16.1)	
	$C_{elevation}$	22.8	-8610	-0.154×10^{-1}	22.7	
		(0.689)	(-3.12)	(-0.049)	(8.56)	
	$C_{riparian}$	0.502×10^{-1}	-0.154×10^{-1}	-0.434×10^{-3}	-0.911×10^{-3}	
		(8.95)	(-0.049)	(-5.13)	(-1.83)	
	$C_{isolation}$	0.652	22.7	-0.911×10^{-3}	-0.43	
		(16.1)	(8.56)	(-1.83)	(-60.8)	
	observations	4778				
	corrected r^2	0.135	0.054	0.242	0.505	
	Wald Test on linear restrictions			264		
RUM	v =	basal area	elevation	riparian	isolation	travel cost
	coefficient	0.567	2.43×10^{-3}	42.0	1.62	-2.94×10^{-2}
		(28.0)	(7.93)	(31.1)	(25.3)	(-43.5)
		(basal area)^2	(elevation)^2	(riparian)^2	(isolation)^2	basal area*riparian
	coefficient	-1.95×10^{-3}	2.08×10^{-7}	-13.5	1.69×10^{-2}	-0.346
		(-19.6)	(7.63)	(-27.4)	(9.30)	(-31.1)
		isol. *riparian	elev.*riparian	elevation*basal	elevation*isol.	basal area*isol.
	coefficient	-0.882	-1.31×10^{-3}	-3.55×10^{-5}	-1.75×10^{-4}	-1.34×10^{-2}
		(-29.1)	(-10.4)	(-18.7)	(-27.2)	(-23.2)
	observations = 4778		percent sites correctly predicted		31.94	

Table IIb: The Parameters of the Inverse Demand Functions: Quadratic Utility
(bold face indicates coefficients are not different at the 5% significance level)

		$C_{basal\ area}$	$C_{elevation}$	$C_{riparian}$	$C_{isolation}$
constant					
	HTC	850	**6.19**	96700	1430
	RUM	19.3	**8.26×10^{-2}**	1430	55.0
	Wald Test	>1000	**4.22×10^{-3}**	>1000	>1000
$\beta_{basal\ area}$					
	HTC	**-3.53**	**-2.39×10^{-2}**	-399	**-5.77**
	RUM	**-0.133**	**-1.21×10^{-3}**	-11.8	**-0.454**
	Wald Test	**.0007**	**2.50×10^{-12}**	>1000	**1.51×10^{-2}**
$\beta_{elevation}$					
	HTC	**-2.39×10^{-2}**	-2.96×10^{-4}	**-2.68**	**-4.62×10^{-2}**
	RUM	**-1.21×10^{-3}**	1.41×10^{-5}	**-4.46×10^{-2}**	**-5.96×10^{-3}**
	Wald Test	**2.50×10^{-12}**	3.34×10^{-19}	**1.36×10^{-4}**	**1.14×10^{-10}**
$\beta_{riparian}$					
	HTC	-399	**-2.68**	-47500	-647
	RUM	-11.8	**-4.46×10^{-2}**	-920	-30.0
	Wald Test	>1000	**1.36×10^{-4}**	>1000	>1000
$\beta_{isolation}$					
	HTC	**-5.77**	**-4.62×10^{-2}**	-647	**-12.2**
	RUM	**-0.454**	**-5.96×10^{-3}**	-30.0	**1.15**
	Wald Test	**1.51×10^{-2}**	**1.14×10^{-10}**	>1000	**2.91**

Table III: Welfare Estimates for A Change in Each Attribute for All Trails (US$/trip)

	Linear Utility				Quadratic Utility			
	10% Decrease in Each Trail Attribute							
	basal area	elevation	riparian	isolation	basal area	elevation	riparian	isolation
RUM	-2.82	2.80	0.30	-0.77	2.72	4.13	1.13	-0.16
HTC	-0.65	-0.97	0.03	0.47	-33.50	- 8.45	-12.50	-1.10
SAH					-24.05	-19.15	-14.90	-2.44

Table III: Welfare Estimates for A Change in Each Attribute for All Trails (US$/trip)

	Linear Utility				Quadratic Utility			
	10% Decrease in Each Trail Attribute							
	basal area	elevation	riparian	isolation	basal area	elevation	riparian	isolation
RUM	-2.82	2.80	0.30	-0.77	2.72	4.13	1.13	-0.16
HTC	-0.65	-0.97	0.03	0.47	-33.50	- 8.45	-12.50	-1.10
SAH					-24.05	-19.15	-14.90	-2.44

Table IV: Welfare Estimates for A Change in Each Attribute of One Trail
(Pleasant Garden Overlook Trail, Unaka Wilderness Area, TN), ($/trip)

	Linear Utility				Quadratic Utility			
	Decrease in Each Trail Attribute of 10%							
	basal area	elevation	riparian	isolation	basal area	elevation	riparian	isolation
RUM	-0.07	0.08	-0.01	-0.02	-0.14	0.01	-0.01	0.07
HTC	-0.01	-0.01	-0.00	-0.00	-3.69	-0.95	-2.20	-0.36
SAH	n/a	n/a	n/a	n/a	-0.42	-0.42	-0.42	-0.42

Appendix A

Summary Statistics

Attribute	Description	Sample Mean (standard deviation)
Basal area	Square feet of trees/acre	65.1 (21.0)
Riparian	% of trail along riparian	34.4 (34.0)
Elevation	Maximum elevation of trail	3330 (1090)
Isolation	Miles from paved road to trailhead	4.45 (4.65)

List of Symbols

ϕ	phi, script in equations 1, 4, 5, 6, and in the text of section 1.1
$\boldsymbol{\phi}$	vector (bold) in rest of text
ε	epsilon, script
∞	infinity, script
λ	lambda, script
α	alpha, script
ψ	psi, script

Footnotes:

---

[1] Recall that we are dealing with one attribute per bundle (e.g. a basket of oranges). The extension to n-attributes (e.g. a fruit basket) requires a line integral but is otherwise the same principle.

[2] The reader is directed to Rosen (1974) and others for discussions of the hedonic methodologies and to McFadden (1978) and others for expositions on the application of the RUM.

Using Multiple-Bounded Questions to Incorporate Preference Uncertainty in Non-market Valuation

Anna Alberini
Department of Economics
University of Colorado, Boulder

Kevin Boyle
Department of Natural Resource Economics and Policy
University of Maine, Orono

and

Michael Welsh
Hagler-Bailly Consulting
Madison, Wisconsin

1. Introduction

The dichotomous choice approach has by now established itself as one of the most popular methods of eliciting information about willingness to pay for a public good or an environmental amenity through a contingent valuation survey. In a typical dichotomous choice contingent valuation survey, respondents are asked whether they would vote in favor or against a proposition on ballot, which, if approved, would provide the commodity at a stated cost to the respondent's household. The "yes" and "no" responses to such cost amounts are analyzed to derive the willingness to pay (WTP) survival curve and the appropriate welfare measures.

The NOAA Panel on Contingent Valuation (1993) sanctioned the use of the dichotomous choice approach and raised the issue of uncovering respondent uncertainty about the value of the commodity. Specifically, the Panel recommended including an explicit "don't know" or "would not vote" response option to the payment question, in addition to "yes" and "no." Unfortunately, it did not offer guidance as to how such responses should be interpreted when modeling WTP.

Several recent applications of the contingent valuation method have indeed modified the traditional dichotomous choice payment question to obtain some measure of the respondent's degree of certainty about willingness to pay. Ready et al. (1995) break down the response to the payment question into six categories (definitely yes, probably yes, maybe yes, maybe no, probably no, and definitely no) and derive ambivalence bounds[1] around WTP for wetland preservation and state-provided incentives to horse farming in Kentucky.[2]

Champ et al. (1997) devise two treatments in their split-sample study about a project removing roads from a portion of the Grand Canyon classified as wilderness area.[3] Respondents in one group were given a polychotomous choice payment question phrased much like that in the Ready et al study, while respondents in the other group were given the traditional dichotomous choice payment question, followed by a question asking them to indicate on a scale from 1 to 10 how certain they felt they would make the payment. Champ et al. find that the actual behavior of respondents is well predicted by hypothetical behavior only for those persons who felt extremely confident about their answers in the survey.

Wang (1997) assumes a respondent refers to a *distribution* of WTP, rather than to a single amount. Only those respondents whose WTP amount is sufficiently large relative to the bid offered in the survey will answer positively to the payment question, while only those respondents whose WTP amount is sufficiently low relative to the bid will decline to pay the bid. All other respondents choose to answer "don't know."

The multiple bounded approach, first proposed by Welsh and Bishop (1993) is yet another approach that allows for varying degrees of uncertainty in the responses to the payment questions. Poe and Welsh (1996) investigate its efficiency properties and compare the distribution of the responses and the welfare measure with those of other elicitation methods. They find that, given their interpretation of the multiple bounded responses, this approach gives statistically efficient estimates and is robust to relatively poor bid designs. They also show that the WTP survival curve based on various interpretations of the multiple-bounded responses can be used to provide bounds for the

WTP curves from the traditional dichomotomous choice, the open-ended and the payment card approaches.

In this paper, we model responses to polychotomous choice payment questions under alternative assumptions about the underlying WTP amount and the respondent's ability to search his or her preferences. Our results show that the point estimates of the welfare measures and the confidence intervals around such estimates vary widely, depending on the statistical model of the data. These findings suggest that when designing a contingent valuation survey that allows for middle responses it is essential to *understand* – rather than guessing ex post – how respondents form their answers and convey uncertainty about their valuation of the resource in question.

The paper is organized as follows. We briefly describe the nature of polychotomous choice/multiple-bounded WTP questions in section 2. The alternative statistical models of WTP compatible with polychotomous choice data are described in section 3. Section 4 presents the data from the Maine ice fishing survey, and estimation results. Section 5 concludes.

2. The multiple-bounded approach.

In a multiple-bounded CV survey, respondents are presented with a *range* of bid values and a number of response categories arranged in a matrix, and are asked to check the degree of confidence with which they feel they would or would not pay *each* amount listed on the card, as shown in Figure 1 below.

In Figure 1 we used five possible response categories, but these could be expanded to include more options, collapsed into fewer options, or labeled using a numerical scale to denote the strength of the respondent's beliefs (Loomis and Ekstrand, 1997).

This question format is widely used in opinion surveys, the goal of the survey often being that of identifying the kind of respondents that feel particularly decisive or undecisive about certain issues. The responses collected in this fashion are usually cross-tabulated against an explanatory variable in a contingency table, and the frequencies in each cell are tested against the hypothesis of independence to determine whether the distribution of the responses differs systematically across groups of respondents (see Agresti, 1996).

In this paper we analyze the determinants of the strength of the responses to multiple bounded payment questions using data collected in a survey of Maine residents about their ice fishing trips, and then attempt to answer the more challenging question of how the responses should be interpreted and statistically model if one is to obtain welfare measures, such as mean and median WTP, as well as confidence intervals around them.

Figure 1.
Multiple-bounded payment questions.

Cost	Definitely Yes	Probably Yes	Not Sure	Probably No	Definitely No
$1	☐	☐	☐	☐	☐
$5	☐	☐	☐	☐	☐
$10	☐	☐	☐	☐	☐
$20	☐	☐	☐	☐	☐
$50	☐	☐	☐	☐	☐
$100	☐	☐	☐	☐	☐
$500	☐	☐	☐	☐	☐
$1000	☐	☐	☐	☐	☐

Because of its similarity to the payment card method, an approach to eliciting WTP that asks respondents to pick a value out of those listed on a card, the multiple-bounded approach may engender the same response biases that the payment card approach has been criticized for (Mitchell and Carson, 1989). Respondents, for instance, may limit their implied WTP amount to the range shown on the card, or tend to "agree" with the middle bid values.[4]

Another important question is whether allowing for uncertain responses ends up actually *encouraging* uncertain responses, i.e., respondents might search their preferences less thoroughly than they would have in the absence of such uncertain response category.

Carson et al. (1995) administer two separate sample of respondents survey instruments that are identical in all respects but the response categories of the vote question: Only one of the two variants of the survey instrument explicitly contains the "would not vote" option. Carson et al. find that when the "would not vote" answer category is explicitly offered to respondents, subjects choose this option more frequently (about 18 percent of the times) than they would spontaneously mention in response to a standard dichotomous choice payment question (8 percent of the times). However, the split between yes and no votes at each level of the bid is not statistically different between the two samples.

3. Statistical Models of Responses

A. The Welsh and Bishop interpretation

Welsh and Bishop (1993) assume that a single, unobserved WTP amount drives all of a subject's responses to the multiple bound payment questions, and develop a statistical framework involving interval data. Specifically, the information coming from all of the responses is collapsed into a single, and relatively tight, interval around the respondent's unobserved WTP value.

To illustrate, consider a person who checks "probably yes" at $5, but "not sure" at $10. One way to interpret this response is to treat the respondent as being willing to pay $5, but not $10. Hence, willingness to pay lies within the interval between $5 and $10, and this person's contribution to the likelihood function is the probability that this subject's WTP amount is bracketed by $5 and $10. The log likelihood function for the sample is:

(1) $$\log L = \sum_{i=1}^{n} \log\left[\Pr(WTP < \$X_i^H) - \Pr(WTP < \$X_i^L)\right]$$

where $\$X_{iL}$ is the highest amount at which respondent i answered "probably yes" and $\$X_{iH}$ is the amount at which the subject switched to a "not sure" response. If willingness to pay is assumed to follow a distribution function F indexed by vector of parameters θ, the method of maximum likelihood can be used to obtain estimates of θ and hence the probability that WTP falls within a specified interval. Poe and Welsh (1996) assume that WTP is distributed as a logistic, but other distributions are possible.

Notice that respondents who never switched to a "not sure" response are treated as if their WTP amount is greater than the highest bid on the payment card, whereas respondents who never answered "definitely" or "probably yes" are assumed to have a WTP amount less than the smallest figure appearing on the card. The contribution to the likelihood is specialized accordingly.

B. Recoding responses

As an alternative interpretation, we posit that respondents revise their underlying valuation of the commodity when answering the multiple bound payment questions. Hence, the response at each dollar amount is motivated by $WTP_{ij} = x_{ij}\beta + \varepsilon_{ij}$, with ε_{ij} independently and identically distributed for all i's and j's, where i indexes the respondent and j indexes the payment amount within the questionnaire. Each revised amount is not directly observed by the researcher.

To develop a statistical model of WTP, the actual responses are recoded following the procedure illustrated by Ready et al. (1995) and the Champ et al. (1997). Specifically, the polychotomous choice responses are reclassified into simple yes/no indicators and traditional single-bounded models of willingness to pay are estimated.

Given our assumptions, this yields an independent probit model that stacks all recoded yes/no responses at all bid levels and for all individuals, and includes the desired independent variables, plus the bid, at the right-hand side. The coefficients of the WTP equation are recovered as $-\gamma_j / \alpha$, j = 1, 2, .., k, where the γ's are the probit estimates of the coefficients of the k independent variables, and α is the coefficient of the bid variable (Cameron and James, 1987).[5]

An important issue here is whether only the "definitely yes" responses should be recoded into simple yes indicators, or probably yes (and maybe even the "not sure" responses) should also be reclassified as a yes.

On comparing the Welsh-Bishop framework with the model(s) based on recoding the responses into simple yes/no indicators, it is immediately apparent that the latter approach artificially inflates the number of observations provided by each respondent, but creates rather broadly defined intervals around the true WTP amount. It is unclear a priori which approach yields more efficient WTP estimates. In addition, recoding into simple yes/no responses effectively neglects the information provided by the respondent as to the degree of uncertainty about their willingness to pay, whereas the Welsh and Bishop interpretation considers the most uncertain responses as the most informative about the true WTP figure.

C. Heteroskedasticity

It is possible that the higher or lower degree of confidence in the responses signals the presence of heteroskedastic error terms ε_{ij} . We argue that, reflecting the difficulty of answering the payment question, the variance of the error term is lower when the bid is sufficiently far away from the person's "true" WTP (i.e., for relatively low or high bids), and higher when the proposed bid amount is very close to the person's "true" WTP.

D. The Random Valuation Model

Our final interpretation of the polychotomous-choice responses is that when the bid is sufficiently close to the expected value of WTP, it becomes more difficult to answer the payment question. Respondents may attempt to search their preferences to resolve such difficulty. Persons with high cost of searching preferences may quit earlier in their searching efforts, and may opt for a "less than positive" answer category.

This raises three related questions. First, can we identify respondents with low and high search costs by examining how the confidence of responses to the payment questions correlates with individual characteristics? Second, can we explicitly incorporate such confidence statements into statistical model of WTP? Third, how do these statistical modeling approaches compare with more traditional procedures?

The random valuation model developed by Wang (1997) is well suited for this interpretation of how subjects answer polychotomous-choice payment questions. We adapt the model to persons with high and low search costs.

Wang (1997) invokes the idea that a respondent considers his entire distribution of WTP when answering. Wang further argues that the individual answers "yes" with probability 1 to a dichotomous choice payment question if the entire WTP distribution lies above the proposed bid amount, and "no" with probability 1 if the entire WTP distribution lies below the proposed bid level. Any other level implies that there is a positive probability that the person agrees to make the proposed payment and a positive probability that he declines. Such a probability is equal to 0.5 if the proposed bid is very close to the mean of the WTP distribution.

In practice, a respondent answers "yes" only if latent WTP amount is sufficiently large relative to the bid, "no" only if latent WTP amount is sufficiently small relative to the bid, and "don't know" if latent WTP amount lies in between. Assuming that WTP is normally distributed, the log likelihood function with three response categories is:

$$\log L = \sum_{i \in yes} \log\left[1 - \Phi\left(\frac{t_i + a_i - x_i\beta}{\sigma}\right)\right] + \sum_{i \in no} \log\left[\Phi\left(\frac{t_i - b_i - x_i\beta}{\sigma}\right)\right] + \tag{2}$$

$$+ \sum_{i \in DK} \log\left[\Phi\left(\frac{t_i + a_i - x_i\beta}{\sigma}\right) - \Phi\left(\frac{t_i - b_i - x_i\beta}{\sigma}\right)\right]$$

where t_i is the cost assigned to the respondent. If a_i and b_i are constants ($a_i \equiv a$ and $b_i \equiv b$ for all i's), then the model is effectively a variant of the ordered probit model (see Greene, 1993). Wang also allows a_i and b_i to be linear functions of a set of individual characteristics: $a_i = z_i\gamma_1$ and $b_i = z_i\gamma_2$. All parameters are identified only if z and x do not include any overlapping variables, or the ratio of a_i to b_i is set to a specified constant.

We adapt the Wang model to the situation with five response modes and independent revisions of the measurement error with which WTP is observed by the respondent. Specifically, we introduce four threshold levels, a, b, c, and d. To keep things manageable, we assume that $c=-b$, and $d=-a$. A respondent answers "definitely yes" to the question if $WTP > Bid + a$, "probably yes" if $Bid + b < WTP < Bid + a$, "not sure" if $Bid - b < WTP < Bid + b$, "probably not" if $Bid - a < WTP < Bid - b$, and "definitely not" if $WTP < Bid - b$. The log likelihood function is:

$$(3) \qquad \log L = \sum_{i=1}^{n} \left\{ \sum_{j \in def.yess} \log\left[1 - \Phi\left(\frac{t_j + a_i - x_i\beta}{\sigma}\right)\right] + \sum_{j \in prob.yes} \log\left[\Phi\left(\frac{t_j + a_i - x_i\beta}{\sigma}\right) - \Phi\left(\frac{t_j + b_i - x_i\beta}{\sigma}\right)\right] + \right.$$

$$\sum_{j \in notsure} \log\left[\Phi\left(\frac{t_j + b_i - x_i\beta}{\sigma}\right) - \Phi\left(\frac{t_j - b_i - x_i\beta}{\sigma}\right)\right] +$$

$$\left. \sum_{j \in prob.no} \log\left[\Phi\left(\frac{t_j - b_i - x_i\beta}{\sigma}\right)\right] + \sum_{j \in DK} \log\left[\Phi\left(\frac{t_j - a_i - x_i\beta}{\sigma}\right)\right]\right\}$$

where observations are independent within a subject and between subjects.

In another variant of equation (3) we allow a, b, c, and d to vary with the respondent, and to be a function of respondent characteristics, mirroring the notion that different people incur different search costs and experience different ability to resolve the difficulty of answering the payment questions.

4. Application to the Maine Ice Fishing Survey

A. The Data

We apply our alternative models to the data collected through a mail survey of Maine residents selected on the basis of fishing license sales records. The survey was conducted after the end of the 1993/1994 fishing season.

Respondents were first asked to provide some general information on their fishing and other consumptive-use activities, and then to answer questions on access to ice fishing sites and contacts with Maine game wardens. Information was obtained about the number of ice fishing trips and ice fishing days experienced in the 1993/94 season, the water bodies visited for ice fishing purposes, type and number of fish caught, expenses incurred, and the respondent's knowledge and support of current fishing regulations and proposed changes. The survey finally inquired about landlocked salmon and cusk fishing.[6]

Right after eliciting the expenses incurred by the respondent during the fishing season, respondents were told: "We would like to know whether you would have gone ice fishing in Maine during the 1993/94 season if your expenditure were <u>more</u> than the total you just reported in Question 4. Please tell us if you would have gone fishing <u>at all</u> at <u>each</u> of the increased costs listed below. (<u>Definitely yes</u> means 'I would have still gone fishing at least once.' <u>Definitely no</u> means 'I would not have gone fishing at all.'). It is very important that you respond to all dollar amounts."

The amounts listed immediately following this question were \$1, \$5, \$10, \$25, \$50, \$75, \$100, \$200, \$300, \$400, \$500, \$1,000, \$1,500 and \$2,000. The response categories were "definitely yes," "probably yes," "not sure," "probably not," and "definitely not."

This question is correctly interpreted as a query on contingent behavior meant to elicit the surplus associated with the current number of fishing trips. Accordingly, we propose and estimate models of the additional amount, Y*, before the respondent's choke price is reached, but our techniques and results are readily extended to more traditional WTP survey and analyses.

B. Distribution of Responses

Descriptive statistics for the sample after records with item non-response were dropped are shown in table 1. Among the individual characteristics, we conjecture that the choke price is likely to depend on respondent income, age and educational attainment. Proxies for fishing experience and commitment to fishing (such as dummies for land-locked salmon, cusk and open water fishing) may capture the cost of searching preferences, and hence influence the likelihood of opting for "not sure" responses.

The majority of the respondents (over 70 percent) checked the "not sure" response option at least once. In fact, most respondents checked *each* response category at least once. Only 17 respondents (0.9 percent of the sample) always answered "definitely yes," and 14 respondents (0.7 percent of the sample) always answered "definitely not."

Table 1.
Descriptive Statistics from the Maine Ice-fishing Survey.

Variable	Mean	Standard Deviation
AGE (years)	38.55	11.97
INCOME (dollars)	38,855	20,961
COLLEGE (dummy variable)	0.90	0.29
MALE (dummy variable)	0.90	0.29
OWFISH (dummy variable for open water fishing)	0.93	0.25
SALMON (dummy variable for landlocked salmon fishing)	0.78	0.41
CUSK (dummy variable for cusk fishing)	0.34	0.47

Figure 2 shows the distribution of responses by bid amount in the Maine ice fishing survey. The percentage of respondents that gave "definitely yes" responses is very high (almost 95%) at very low bid levels and declines sharply as the bid level increases. When the additional cost per season is $2000, only 1 percent of the study participants would definitely be prepared to continue fishing. The percentage of "definitely no" responses is very low at the low bid values, and rises in a regular fashion with the bid. Over ninety-one percent of the respondents would definitely not continue fishing at an additional cost of $2000. The likelihood of providing "definitely" or "not sure" responses is generally low at the lowest and highest bid amounts, and peaks at the middle amounts.

The pattern of response shown in Figure 2 is consistent with the possibility that the highest uncertainty occurs at bid levels close to the person's true surplus, but could also signal the presence of range bias, with respondents tending to switch from positive to negative responses at the central bid amounts. However, we are unable to test for the presence of range bias using these data. A Pearson chi square test easily rejects the null hypothesis that the responses are independent of the bid level at less than the 1 percent level of significance.[7]

C. Predictors of Uncertain Responses

To help identify variables that capture search efforts and costs, we fit a multinomial logit model in which the bid levels listed on the card and individual characteristics are entered as predictors of the response category checked by the respondent at each bid level. Formally, the log likelihood function for the multinomial logit model is:

$$\log L = \sum_{i=1}^{n} \sum_{k=1}^{K} \sum_{j=1}^{J} I_{ikj} \cdot \log \pi_{ikj} \tag{4}$$

where I is a dummy variable indicating whether respondent i selected the j-th response category at the k-th valuation task, and π is the probability that such a selection is made. The probability depends on individual characteristics and the bid value (summarized into the vector z) via a set of choice-specific coefficients:

$$\pi_{ijk} = \frac{\exp(z_{ijk}\beta_j)}{\sum_{l=1}^{J} \exp(z_{ilk}\beta_l)} \quad . \tag{5}$$

The estimates of the choice-specific coefficients β are reported in Table 2 along with their asymptotic t statistics. A positive coefficient indicates that an increase in the value of the independent variable makes the respondent more likely to choose the indicated response category.

Table 2.
Multinomial logit model of response choice.
Omitted category: definitely not. T statistics in parentheses.

	Definitely yes	Probably yes	Not sure	Probably not
Constant	1.10320 (8.03)	-0.05366 (-0.32)	-0.85451 (-5.03)	-1.46815 (-8.90)
college education	0.12530 (1.77)	0.15441 (1.77)	0.02552 (0.30)	0.21372 (2.43)
landlocked salmon	0.52264 (10.42)	0.34584 (5.70)	0.49055 (7.59)	0.27704 (4.53)
open water fishing	0.19925 (2.43)	0.11908 (1.21)	0.05717 (0.57)	0.15530 (1.58)
cusk fishing	0.45039 (9.90)	0.25763 (4.76)	0.22829 (4.22)	0.10381 (1.99)
age	-0.01775 (-10.19)	-0.02055 (-9.55)	-0.01484 (-6.88)	-0.00529 (-2.58)
male	0.32683 (4.64)	0.19988 (2.25)	0.27678 (3.09)	0.06343 (0.78)
income	0.00002 (18.45)	0.00001 (11.48)	0.00001 (6.46)	0.00001 (7.57)
bid	-0.01257 (-68.41)	-0.00559 (-38.83)	-0.00242 (-31.86)	-0.00125 (-26.36)
log likelihood	-26,526.7228			

The results imply that many of the socio-economic and fishing experience variables are significant predictors of response choice. Even more importantly, the bid level is related to response choice.

Older respondents are more likely to select the "definitely not" response categories than any other. Holding all else unchanged, wealthier respondent are more likely to answer "definitely yes." Anglers with cusk fishing experience are more likely to answer "definitely yes" and to steer away from "not sure" or "negative responses." Persons who enjoy land-locked salmon fishing are more likely to answer "definitely yes," "probably yes," "not sure" and "probably not" than they are to answer "definitely no," all else unchanged. Raising the bid would progressively move respondents across the four response categories listed in table 2.

We use likelihood ratio tests to determine which factors are significantly associated with the choice of each response category. The likelihood ratio tests check that all of the four coefficients associated with one variable are significantly different than zero. Only the dummy for open water fishing activities is not significant at the conventional levels, while the dummy for college-level education is significant at the 10 percent level, but not at the 5 percent level.

*D. Statistical Models of Y**

In tables 3 and 4 we report four alternative models of Y*, the unobserved variable denoting how much higher the cost of fishing can go until the respondent forgoes ice fishing altogether. We assume that Y* is normally distributed, and we are interested in recovering mean Y*.[8]

The models we fit include: (1) Welsh-Bishop models under alternative assumptions about which response category (probably yes, not sure, probably not) defines the upper bound of Y*; (2) probit models based on recoding the responses into simple yes and no answers under alternative recoding conventions; (3) models allowing for "definitely," "probably" and "not sure" responses to be driven by error terms with different variances; and (4) models in which the confidence in the commitment to pay depends on the distance between the bid and the center of the respondent's distribution of Y*, with and without allowing for certain individual characteristics and behaviors to capture preference search costs.

In the models with heteroskedasticity, we assume that "not sure" responses signals the point of indifference between the bid and the respondent Y*, and interpret the amounts at which a "not sure" response is given as point estimates of the person's WTP. The likelihood function is, therefore:

(6)

$$\sum_i \sum_j \sum_k \left[\sum_l Not_{ijk} \cdot \log \Phi\left(\frac{t_k - x_i\beta}{\sigma_l} \right) + Yes_{ijk} \cdot \log\left(1 - \Phi\left(\frac{t_k - x_i\beta}{\sigma_l} \right)\right) + NotSure_{ik} \cdot \log \phi\left(\frac{t_k - x_i\beta}{\sigma_{NS}} \right) \right]$$

where i denotes the individual, k denotes the bid listed on the card, l denotes the "definitely" or "probably" response category, and Yes, Not and NotSure are dummy indicators.

One variant of the random valuation model lets $a=-d$ and $b=-c$ of equation (3) be constants. The next variant lets a and b be linear combinations of individual characteristics that our mulinomial logit analysis identified as related to the tendency to answer "not sure," and hence to the search costs. Specifically, we assume $a=z\gamma_1$ and $b=z\gamma_2$ where z includes dummies for whether the respondent goes fishing in open waters and for landlocked salmon. These dummies are excluded from the independent variables x entering in the determination of E(Y*) to ensure identification of all parameters.

The results show that the estimates of E(Y*) can change dramatically from one statistical model to the next and with the alternative recoding of "not sure" responses.

The probit model that restrictively interprets only "definitely yes" responses as true yes responses result in an estimate of mean Y* equal to $98.78. Allowing "probably yes" responses to be interpreted as true yes responses raises mean Y* to about $210 – an increase of over 100 percent. When the "not sure" also also treated as yes, mean Y* rises to about $350. Deleting the "not sure" responses from the usable sample, a common practice in the analysis of contingent valuation survey data, brings down mean Y* to about $131 and $258, respectively.

Interestingly, the standard error around mean Y* changes quite a bit with these alternative recoding convention: the standard error is lowest ($2.53) when only the definitely yes responses are treated as true yes, and *increases* as other response categories are interpreted as true yes. While the split between zeros and ones is made more even (which should make the estimates more efficient and decrease the standard errors), at the same time the distribution of Y* is flattened out, which increases the underlying dispersion of Y*, and hence the standard errors of the estimates of mean Y*.

The Welsh-Bishop models give estimates of mean Y* that are within about 10 percent of the corresponding independent probit estimates. It is surprising that the standard errors around mean Y* are almost or over twice as large as those from the probit models.

Allowing for three separate variances of the error term (with "definitely" responses, whether yes or no, sharing the same variance; "probably" responses, whether yes or no, sharing equal variance; and "not sure" responses being imputed their own) results in mean Y* of $250. The standard error around this estimate is $2.86. Estimated mean Y* jumps to \$388 when five different variances are allowed. The standard error around this estimate is \$5.15.

Finally, the random valuation model with constant thresholds *a* and *b* common to all respondents yields an estimate of mean Y* equal to \$315.77, and implies that the respondent true Y* must be greater than \$248 before a "definitely yes" answer is given, and greater than \$76 before a "probably yes" answer is given.

Table 4 reports results of a subset of the same models, but with E(Y*) expressed as a linear function of individual characteristics. Clearly, the coefficients of the independent variables can vary in magnitude as well as in sign and significance as we move from one model to the next, suggesting that testing hypotheses about how mean Y* is influenced by individual characteristics can be expected to produce widely different conclusions, depending on the model adopted by the researcher.

The two random valuation models displayed in columns 4 and 5 of table 4 differ only in that *a* and *b* are held constant across respondents or allowed to be determined by certain respondent fishing behaviors. The coefficients of the independent variables entering in E(Y*) are very close across these two specifications. T statistics for the γs show that landlocked salmon fishing raises the thresholds above which the respondent answers "definitely yes" and "probably yes," and lowers the thresholds below which the respondent gives a "definitely not" answer, making the uncertainty ranges somewhat broader. This result is consistent with what shown by the multinomial logit regression.

5. Conclusions

We have estimated the mean of the latent variable, Y*, driving responses to polychotomous choice payment questions using different statistical models. These models rest on alternative assumptions about the way respondents form their answers and/or on alternative reclassifications of such answers.

We have found that estimated mean Y* varies widely with the model and the recoding convention. The lowest mean Y* is about \$99, the largest is \$388 – a difference of almost 400 percent. The standard errors around the estimated mean Y* also vary quite a bit, ranging from \$2.53 to \$10.19.

Finally, models of Y* including covariates show that estimated coefficients can vary dramatically in size, sign and significance levels as we move from one model to the next.

The sensitivity of the estimates to the specification of the model suggests that it may be necessary to explore – using focus groups and personal interviews – how respondents react to polychotomous choice questions and how underlying values can be inferred from their responses to gain a better understanding of the appropriate statistical framework for polychotomous choice/multiple bounded responses.

Table 3
Probit Models with recoding

	Def. yes = 1 all else = 0	def. yes = 1 prob.yes = 1 all else = 0	def. yes = 1 prob. yes =1 not sure =1 all else = 0	delete not sure def.yes=1 all else=0	delete not sure def.yes=1 prob.yes=1 all else=0
mean Y*	98.78	209.98	353.53	131.408	257.87
s.e. (mean Y*)	2.53	2.73	4.19	2.62	2.80
Welsh-Bishop models					
	switch away from def. yes	switch away from prob. yes	switch away from not sure		
mean Y*	109.02	192.75	319.46		
s.e. (mean Y*)	5.63	7.12	10.19		
Heteroskedasticity models					
	three variances	five variances			
mean Y*	250.05	388.15			
s.e. (mean Y*)	2.86	5.15			
Random Valuation models (A,B constant) (t statistics in parentheses)					
mean Y*	315.77	γ1	247.93 (78.31)		
s.e. (mean Y*)	3.85	γ2	76.24 (46.00)		

Table 4
(t statistics in parentheses)

	Independent Probit	Welsh-Bishop	Hetero-skedasticity	Random Valuation	Random Valuation
	Def.yes=1 prob.yes=1 all else=0	Switch away from prob.yes	three variances	*a* and *b* are constants	A, B linear functions of OWFISH and SALMON
Constant	104.24 (5.98)	84.99 (1.82)	140.84 (8.101)	219.57 (10.41)	219.7266 (10.39)
OWFISH	22.76 (2.14)	35.77 (1.28)	24.18 (2.303)		
SALMON	39.73 (6.22)	28.18 (1.63)	52.52 (8.19)		
CUSK	51.412 (8.99)	60.36 (4.05)	53.72 (9.53)		
AGE	2.18 (9.68)	-1.87 (-1.72)	-2.23 (-10.12)	-3.05 (-9.48)	-3.03 (-9.43)
INCOME	0.002 (15.64)	0.002 (5.42)	0.0020 (15.05)	0.0032 (17.03)	0.0032 (17.09)
MALE	35.46 (3.81)	46.50 (1.93)	43.19 (4.77)	65.72 (5.07)	65.99 (5.06)
COLLEGE	11.96 (1.37)	-5.64 (-0.23)	2.99 (11.64)	44.20 (3.38)	43.22 (3.29)
σ	294.36 (215.93)	300.15 (57.91)	234.23 (72.385) (def.yes, def.no) 513.94 (25.695) (prob yes, prob no) 354.53 (65.149) (not sure)	493.74 (93.60)	493.28 (93.62)
				A=248.16 (78.824) B= 76.35 (46.11)	γ10=227.11 (41.96) γ11=-0.0195 (-0.19) γ12=26.88 (4.62) γ20=65.78 (11.94) γ21=-3.08 (-0.61)

REFERENCES

Agresti, Alan (1996), *An Introduction to Categorical Data Analysis*, New York: Wiley & Sons.

Alberini, Anna (1995), "Optimal Designs for Discrete Choice Contingent Valuation Surveys: Single-bound, Double-bound and Bivariate Models," *Journal of Environmental Economics and Management*, 28, 187-306.

Arrow, Kenneth, Robert Solow, Paul R. Portney, Edward E. Leamer, Roy Radner, and Howard Schuman (1993), "Report of the NOAA Panel on Contingent Valuation," *Federal Register* 58, 4601-4614.

Champ, Patricia A. , Richard C. Bishop, Thomas C. Brown and Daniel W. McCollum (1997), "Using Donation Mechanisms to value Nonuse Benefits from Public Goods," *Journal of Environmental Economics and Management*, 33, 151-162.

Cameron, Trudy A. (1991), "Interval Estimates of Non-market Resource Values from Referendum Contingent Valuation Surveys," *Land Economics*, 67, 413-421.

Cameron, Trudy A. and Michelle James (1987), "Efficient Estimation Methods for 'Closed-ended' Contingent Valuation Surveys," *The Review of Economics and Statistics*, 69, 269-276.

Carson, Richard T., W. Michael Hanemann, Raymond J. Kopp, Jon A. Krosnick, Robert C. Mitchell, Stanley Presser, Paul A. Ruud, and V. Kerry Smith (1995), "Referendum Design and Contingent Valuation: The NOAA Panel's No-Vote Recommendation," Resources for the Future Discussion paper QE96-05, Washington, DC, November.

Hanemann, W. Michael (1984), "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses," *American Journal of Agricultural Economics*, 66, 332-341.

Kanninen, Barbara J. (1991), "Optimal Experimental Design for Contingent Valuation Surveys," Ph.D. dissertation, University of California Berkeley.

Loomis, John B. and Earl Ekstrand (1997), "Alternative Approaches for Incorporating Respondent Uncertainty when Estimating Willingness to Pay: The Case of the Mexican Spotted Owl," Dept. of Agricultural and Resource Economics, Colorado State University, Fort Collins, CO, April.

Mitchell, Robert C. and Richard T. Carson (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Washington, DC: Resources for the Future.

Park, Timothy A., John B. Loomis and Michael Creel (1991), "Confidence Intervals for Evaluating Benefit Estimates from Dichotomous Choice Contingent Valuation Studies," *Land Economics*, 67, 64-73.

Poe, Gregory L. and Michael P. Welsh (1996), "WTP Certainty Intervals and the Disparity Between Alternative Contingent Valuation Elicitation Formats: Application of a Multiple Discrete Choice Approach," Cornell University Working Paper Series in Environmental and Resource Economics 96-03, June.

Ready, Richard C., John C. Whitehead and Glenn C. Blomquist (1995), "Contingent Valuation When Respondents Are Ambivalent," *Journal of Environmental Economics and Management*, 29, 181-196.

Rowe, Robert D., William D. Schulze and William S. Breffle (1996), "A Test for Payment Card Biases," *Journal of Environmental Economics and Management*, 31, 178-185.

Wang, Hua (1997), "Treatment of 'Don't Know' Responses in Contingent Valuation Survey: A Random Valuation Model," *Journal of Environmental Economics and Management*, 32, 219-232.

Welsh, Michael P. and Richard C. Bishop (1993), "Multiple Bounded Discrete Choice Models," in *Benefits and Costs Transfer in Natural Resource Planning*, Western Regional Research Publication, W-133, Interim Report 6.

Footnotes:

[1] In an ambivalence region, trading off income for the commodity is not clearly superior or inferior to the status quo.

[2] The ambivalence bounds, defined as the lowest amount at which 50 percent of the respondents give a "definitely no" answer and the highest amount at which 50 percent of the respondents give a "definitely yes" response, are shown to be quite large. For the scenarios involving wetlands preservation, the tightest ambivalence interval spans between a few pennies and $20, while the broadest spans between a few pennies and $157.75.

[3] This study is different than most contingent valuation surveys because it solicits donations to a relatively small and low cost project.

[4] Rowe et al (1996), however, have recently questioned the notion that the payment card truly biases responses. A comparison of four independent samples of subjects that were given payment cards reporting different ranges of values shows that WTP does not significantly vary with the variant of the payment card, as long as the payment card does not truncate the upper end of the value distribution.

[5] Various methods have been proposed to obtain standard errors around the estimated coefficients, β, and mean WTP. Possible alternatives includes the fiducial approach (adapted to discrete choice contingent valuation data by Kanninen, 1991, and Alberini, 1995), use of first-order Taylor series expansion approximations (applied by Cameron, 1991), and bootstrapping techniques (Park, Loomis and Creel, 1991).

[6] Cusk an alternative name for the burbot, a freshwater fish.

[7] If the choice of a response category is truly independent of the bid level, the frequencies along the rows of a contingency table crossing the bid levels against the response categories should remain approximately the same. The test statistic is $\chi^2 = \sum (n_{ij} - \mu_{ij})^2 / \mu_{ij}$, where n_{ij} are the observed frequencies, and μ_{ij} are the frequencies predicted by the independence model (Agresti, 1996). In this particular case, since all of the n respondents are confronted with the complete list of payment levels, $\mu_{ij} = n\pi_{+j}$, where π_{+j} is the marginal probability of each response category. Here, the test is distributed a chi square with 52 degrees of freedom. The chi square was computed to be 19,935.63, which falls in the rejection rejection of the chi square with 52 degrees of freedom at conventional levels of significance.

[8] For the sake of simplicity, we report the regular estimate of mean surplus, rather than using the formula (Hanemann, 1984) that truncates such measure at zero, as is frequently done in practice.

Valuing Stated Preferences For Health Benefits of Improved Air Quality: Results of a Pilot Study

William H. Desvousges, F. Reed Johnson,
Melissa C. Ruby, and Alicia R. Gable

Triangle Economic Research
1000 Park Forty Plaza, Suite 200
Durham, NC 27713
Phone: (919) 544-2244 Fax: (919) 544-3935
Email: info@ter.com Web: www.ter.com

Abstract

This paper presents a methodology and empirical estimates from a pilot study for estimating willingness to pay for health improvements associated with reduced exposure to air pollution. The pilot study uses a rated-pair format to elicit stated preferences for various health-state attributes and costs. This approach has the advantage over other valuation approaches in that it is utility-theoretic and can elicit WTP values for a variety of health outcomes from both symptomatics and nonsymptomatics. Illustrative WTP results exemplify the flexibility and potential of this valuation approach for estimating the benefits of health improvements for a variety of policy purposes.

INTRODUCTION

This paper presents initial results from a cooperative effort between Triangle Economic Research (TER), Health Canada, Environment Canada, Ontario Hydro, Ontario Ministry of Environment and Energy, and the Quebec Ministry of Environment. The primary objective of this effort was to design, prepare, and test a survey instrument to measure Canadian willingness to pay to reduce the morbidity effects of air pollution. This paper summarizes the development of the survey instrument, provides the results of the pilot test used to assess the survey instrument, and discusses recommendations for the administration of the full-scale survey.

Regulatory programs, including air quality regulations, often are intended to improve public health. Measuring the benefits of health improvements is a challenging endeavor because of the many different parties involved, including individuals with varying susceptibilities to ill health, health-care providers, third-party insurers, and society in general; and because of the different types of benefits to be measured, including individual benefits and collective benefits. This study estimates individuals' willingness to pay (WTP), that is, the sum of any actual expenditures and consumer surplus, for an incremental improvement in health. Health effects include episodes of mild to severe respiratory and cardiac illnesses that epidemiological studies have linked to air pollution. An *ex post* perspective is adopted in the valuation for both practical and conceptual reasons. This decision implies that the study is valuing the reduction of an episode of a given health outcome, not the risks of experiencing the outcome.

Although markets exist for some aspects of health, the existence of third-party payments alters the relationship between supply and demand. Thus simple market analysis is not sufficient for estimating values for health. Most previous health studies have been conducted using cost-of-illness or contingent-valuation methods. However, because of the serious limitation associated with these approaches, stated-preference analysis is used to elicit WTP values.

Stated-preference (SP) experiments recognize that commodities have value because of their attributes. SP experiments have been used extensively in marketing research and product development (Cattin and Wittink, 1982; Wittink and Cattin, 1989). Specific marketing applications have been aimed at new-product identification, market segmentation, advertising, distribution, competitive analysis, and price optimization. In recent years, SP has been applied in the field of environmental economics as an alternative to the CV method. Two recent studies, Viscusi, Magat, and Huber (1991) and Krupnick and Cropper (1992) (using the Viscusi data), use SP to elicit a value from respondents for reducing chronic health risks.[1]

Preferences for health states are analogous to other commodities within an SP framework. Each health state is made up of several attributes. For example, in our study, the attributes of the health condition included the number of episodes, the symptoms, the level of daily-activity functioning, and a cost attribute. We presume that people have preferences for different levels within these attributes and are willing to accept some trade-offs among them. Their preferences for different health states are indicated by the revealed trade-offs. For example, people may be willing to trade some limitations in daily activity for decreased episodes. SP is designed to measure the rates at which people are willing to accept such trade-offs. By including a monetary cost, we can express these trade-offs in dollar terms, or WTP.

ELICITING STATED PREFERENCES USING RATED PAIRS

Three types of SP experiments lend themselves to the valuation of health effects: rated pairs, discrete choice, and ranking. This study employs the rated-pair format,

---

[1] The resource-economics literature also is beginning to see some applications of SP. Gan and Luzar (1993) use SP to value hunting trips in Louisiana. MacKenzie (1993) values hunting trips in Delaware using SP analysis. Opaluch et al. (1993) also use SP to describe public preferences for siting a noxious facility. Adamowicz, Louviere, and Williams (1994) use SP to explain recreational site choice selection. Johnson et al. (1995) use SP to estimate electric customers' willingness to pay for environmental and other attributes of electricity generation. Roe, Boyle, and Teisl (1994) use SP to value the effects on sport fishing of implementing alternative management plans to restore runs of Atlantic salmon in Maine.

which measures respondents' valuations of slight variations in attributes by requiring them to evaluate trade-offs among various attributes. Respondents are sequentially presented with several different pairs of bundled commodities, represented as sets of attribute levels, and asked to compare each pair. They are asked to rate the intensity of their preference for one of the pairs on a numerical scale, such as 1 to 7, where 1 indicates a strong preference for the first program, 7 indicates a strong preference for the second program, and 4 indicates indifference between the two programs. The respondent is asked to rate a series of these pairs, with each pair having different attributes or attribute levels.[2]

Figure 1 shows an example of a rated-pair screen used in this study. In this example, the price is expressed as health-maintenance costs.[3] The respondents indicate their preferences for Condition A versus Condition B. The complete SP exercise presents a series of pairs to respondents and records their ratings. These stated preferences serve as the data needed for estimating the underlying health-state utility function for these attributes. By including price as one of the attributes, it is possible to rescale the utility index to dollars and derive estimates of willingness to pay for different health states and their attributes.[4]

---

[2] This approach was used in the Viscusi, Magat, and Huber (1991) study valuing bronchitis risks.

[3] See the section of this paper, "Development and Pilot-Test Administration of the Stated-Preference Survey Instrument" (beginning at the bottom of p. 8) for a discussion of the payment vehicle.

[4] All dollars expressed in this paper are Canadian dollars.

Figure 1.
Example of Rated-Pair Stated Preference Question

Category	Condition A	Condition B
Number of episodes this year	3 episodes lasting 7 days	4 episodes lasting 7 days
Symptom	Coughing, wheezing, shortness of breath	Coughing, wheezing, shortness of breath
Daily Activities	• CANNOT leave your house, go to work, go to school, do housework or participate in social or recreational activities • Have SOME physical limitations • CAN care for yourself	• CANNOT leave your house, go to work, go to school, do housework or participate in social or recreational activities • Are in hospital • Need help caring for yourself
Costs	Total costs of $700 this year to your household	Total costs of $200 this year to your household

1	2	3	4	5	6	7
A is much better	A is somewhat better	A is slightly better	A and B are about equal	B is slightly better	B is somewhat better	B is much better

ESTIMATING WILLINGNESS TO PAY FOR CHANGES IN HEALTH STATES

A primary study objective is to estimate the effects of changes in health-state attributes and costs on respondents' well-being as indicated by their ratings of the SP profiles. Thus, the analysis goal is to estimate a function that maps attributes and costs into a utility index that is consistent with the observed rating data.

We assume that individual indirect utility can be expressed as a function of commodity attributes and personal characteristics:

$$U_t^i = V^i(X_t, Z^i, p_t; \beta, \gamma, \delta) + e_t^i \qquad (1)$$

where

U_t^i is individual i's utility for attribute profile t,

$V^i(\cdot)$ is the non-stochastic part of the utility function,

X_t is a vector of attribute levels in individual i's choice set,

Z^i is a vector of personal characteristics,

p_t is the cost of the commodity bundle,

e_t^i is a disturbance term,

β is a vector of attribute parameters,

γ is a vector of individual-specific parameters, and

δ is the cost parameter.

The attribute, individual-specific, and cost parameters are estimates of the marginal effects on utility of attributes, individual tastes, and money.

Let tR and tL denote the right-side and left-side commodity profiles for profile pair t, respectively. The utility difference for commodity pair t is simply

$$dU_t^i = V_{tR}^i - V_{tL}^i + \varepsilon_t^i \qquad (2)$$

where dU^i_t is the difference in respondent i's utility for profile pair t. V_{tR}^i and V_{tL}^i are the indirect utilities associated with the right-side and left-side profiles, respectively, and $\varepsilon^i_t = e^i_{tR} - e^i_{tL}$ is the associated disturbance term. The disturbance term captures the effects of unobserved factors, including possible inherent ambiguity of respondent preferences and cognitive errors.

The difference in indirect utility for commodity pair t, dU^i_t, is specified as a linear function of health-state attributes and the log of cost, as shown in Equations (3) and (4).[5]

$$dU_t^i = V_{tR} - V_{tL} + \varepsilon_t^i = \left[\sum_j \beta_j \cdot X_{jtR} + \delta^i \cdot \ln(COST_{tR})\right] - \left[\sum_j \beta_j \cdot X_{jtL} + \delta^i \cdot \ln(COST_{tL})\right] + \varepsilon_t \quad (3)$$

---

[5] By using the log of cost, the marginal utility of money (δ^i) is allowed to vary in a nonlinear way across costs. Thus the marginal utility of money is $\frac{d[\delta^i \ln(COST)]}{dCOST} = \frac{\delta^i}{COST}$ instead of δ^i.

$$\delta^i = \gamma_o + \sum_k \gamma_k \cdot Z_k^i \qquad (4)$$

where j represents one of the attributes in the attribute bundle.[6] The marginal utility of money provides a means of scaling changes in the health-state utility index in dollar equivalents. We allow the marginal utility of money, δ^i, to depend on personal characteristics, which allows heterogeneous tastes to affect the relative utilities of health and money, even though the β_j parameters are constant across respondents.

This specification assumes that attributes neither are substitutes nor complements for each other, so a change in the level of one attribute does not affect the marginal utility of any other attribute.[7] It also assumes that all respondents share common utility-function parameters (β). However, because SP surveys collect multiple responses for each person (12 morbidity ratings in our design), it is more appropriate statistically to estimate a panel model that controls for respondent differences.

Utility difference dU^i_t in Equations (2) and (3) is not directly observable. Instead, we observe R^i_t, which is a discrete rating category related to the unobserved dU^i_t of interest. Thus Roe, Boyle, and Teisl (1994) argue that dummy regression estimation of attribute marginal utilities is not appropriate for computing welfare measures. Furthermore, multinomial logit estimation fails to take into account the ordinal scale of the response categories. The appropriate approach, therefore, is ordered logit or probit, which incorporate both the discreteness and the natural ordering of the data.

This paper used ordered probit which assumes the unobserved error term is normally distributed. To estimate the ordered probit models, the data are sorted so that

---

[6] If the attribute levels are continuous, estimation can be simplified by treating the utility difference as the utility of the difference in attribute levels. This specification permits the use of commonly available statistical-estimation software. Unfortunately, both symptoms and activity limitations are discrete health states, so it was not possible to implement this simplification for this study.

[7] See Keeney and Raiffa (1978) for an analysis of the properties of such utility functions.

the preferred profile is on the right, making $dV^i_t = V^i_{tR} - V^i_{tL} \geq 0$.[8] We construct the ranking categories, R^i_t, by recoding responses accordingly, so that zero indicates indifference and three indicates maximum difference.[9]

Because probit assumes the Equation (3) error term $\varepsilon^i_t \sim N(0,\sigma^2)$, the probability of observing response R^i_t is

$$\Pr ob(R^i_t = k) = \Phi\left\{\frac{\alpha_k - dV^i_t}{\sigma}\right\} - \Phi\left\{\frac{\alpha_{k-1} - dV^i_t}{\sigma}\right\} \quad k = 0,1,..,3 \tag{5}$$

where Φ is the cumulative standard normal distribution function.[10] Scaling the difference between α_k and dV by the standard deviation σ enables us to exploit the known properties of the standard normal distribution. The maximum-likelihood estimation procedure estimates threshold and utility parameters that yield probabilities that correspond to the observed proportions of responses in the various rating categories.[11]

Estimating the parameters of the utility function enables us to quantify the value of changes in health state. The marginal utility of money is the increased number of utility units corresponding to a one-dollar increase in purchasing power. Thus any change in utility induced by a change in health state can be converted to its dollar equivalence by dividing it by the marginal utility of money.

---

[8] This procedure assumes that respondents have no systematic preference for screen location.

[9] Because the original response scale indicates both which profile is preferred and how much it is preferred, this rearrangement maps response 3 into 5, 2 into 6, and 1 into 7. Response 7 indicates maximum utility difference and 4 indicates indifference, so R^i_t equals the recoded response minus 4.

[10] The maximum-likelihood procedure used to estimate the model parameters normalizes the α_{-1} threshold at $-\infty$ and α_3 at $+\infty$ and does not include an intercept term.

[11] In principle, the α thresholds also may vary across individuals, as in a fixed-effects model. However, this requires computing a T-fold multiple integral, where T is the length of the time series. This is computationally infeasible for T greater than 4 or 5. However, simulation methods are available for solving such problems. (See, for example, Train, 1995.)

In the specification of Equation (3), the β coefficient for each attribute represents its constant marginal utility. The δ coefficient of the price attribute is interpreted as the marginal utility of money. The willingness to pay for a given change in health state (X_j - X^*_j) is the amount of money (p^*_j - p_j) that would leave the respondent indifferent between the payment and the change in health.

$$V^i(X_j^*, Z^i, p_j; \beta, \gamma, \delta) = V^i(X_j, Z^i, p_j^*; \beta, \gamma, \delta) \qquad (6)$$

$$WTP_j^i = p_j^* - p_j = \frac{(X_j - X_j^*)\,\partial V^i / \partial X_i}{-\partial V^i / \partial p} = \frac{(X_j - X_j^*)\beta_j}{-\delta} \qquad (7)$$

Any payment less than or equal to WTP^i_j leaves individuals at least as well off as they would be if the change (X_j - X^*_j) had not occurred. We used this procedure to calculate the empirical estimates of WTP, modified as required for models that allow parameters to vary with socioeconomic characteristics.

DEVELOPMENT AND PILOT-TEST ADMINISTRATION OF THE STATED-PREFERENCE SURVEY INSTRUMENT

The objective of this project is to design and evaluate a survey instrument to measure the willingness to pay to avoid adverse health effects from air pollution. This section describes the design, content, and development of the survey instrument, and the relevant details of the pilot test used to evaluate the survey instrument.

The SP survey instrument has four attribute categories: symptom, number of episodes, daily activity level, and cost. With assistance from Health Canada, and with information from the pretesting, the appropriate levels of these attribute categories were determined. In addition, pretests showed that having all four attributes change for each SP pair was harder for respondents. As a result, for each rated pair, the symptom is held constant across the pair and the other three attributes vary. Variation in symptoms occurs across pairs.

Table 1 shows the attributes and attribute levels for the experimental design. As shown, the symptoms used are cardiac and respiratory problems that range from relatively mild to more severe, and are all episodic. The change in the number of episodes is small but policy-relevant, with a reduction of one or two episodes. The costs were chosen to be significant enough that the respondents would consider them, but not so large that they would dominate the trade-offs. These costs were presented as out-of-pocket costs related to reducing the severity and frequency of illnesses that are not covered by the government health system or company insurance plan (e.g., vitamins, medicines, air filters, optional treatments). Finally, the daily activity levels are a modified version of the mobility, physical activity, and social activity descriptors used in the Quality of Well-Being (QWB) health status classification system.[12] These activity levels cover a wide range of effects from no physical limitations to confinement to hospital.

Respondents react to the levels of each attribute as well as the differences in episodes, cost, and activity levels within pairs. The levels, therefore, must be sensible to the respondents in order for them to seriously consider the trade-offs. Thus, the levels of episodes and costs shown to the respondents vary for different symptoms. For instance, the symptom "stuffy/runny nose and sore throat" was only seen with cost levels ranging between $50 and $550, and episode levels ranging between 3 and 5 episodes. Similarly, some restrictions were imposed on the design to ensure that daily activities were credible to respondents. For instance, the symptom "stuffy/runny nose and sore throat" was never seen with Daily Activity Level 6 (confined to hospital). Other similar restrictions exist in the design.

[12] Health-status indexes, such as QWB, are based on the idea that health is affected by both objective factors, such as behavior and motor function, and subjective factors, such as people's ability to fulfill the roles and expectations they have for themselves. The QWB index defines health states in four dimensions: three function states (mobility, physical activity, and social activity) and the most severe symptom/problem complex.

Table 1.
Attribute and Attribute Levels Shown In Morbidity Comparisons

ATTRIBUTE	LEVEL	DESCRIPTION
Symptom	1	Stuffy/runny nose and sore throat
	2	Eye irritation
	3	Generally tired and weak
	4	Fluttering in chest and feeling light-headed
	5	Coughing, wheezing, shortness of breath
	6	Coughing or wheezing with fever, chills, or aching all over
	7	Shortness of breath, and swelling in ankles and feet
	8	Pain in chest or arm
Episodes	1-5	Episodes of one-week duration
Cost	\$50 – \$700	Health maintenance costs not covered by government or insurance
Daily Activity	1	You can work, go to school, do housework, participate in social or recreational activities, and have no physical limitations.
	2	You can work, go to school, do housework, and participate in social or recreational activities, but you have some physical limitations (trouble bending, stooping, or doing vigorous activities) because of this health condition.
	3	You can go to work, go to school, do housework, but you have some physical limitations (trouble bending, stooping, or doing vigorous activities), and cannot participate in social or recreational activities because of this health condition.
	4	You cannot leave your house, go to work, go to school, do housework, participate in social or recreational activities, and you have some physical limitations (trouble bending, stooping, or doing vigorous activities) because of this health condition, but you can care for yourself.
	5	You cannot leave your house, go to work, go to school, do housework, participate in social or recreational activities, and you need help caring for yourself (feeding, bathing, dressing, toilet) because of this health condition.
	6	You are in hospital and need help caring for yourself (feeding, bathing, dressing, toilet).

Even with these restrictions, it is not possible for each respondent to see every possible combination of attributes. In order to limit the length of the survey and to avoid cognitive burden, each respondent saw only 12 pairs of attribute combinations. The computer program randomly drew these 12 pairs from a restricted design space to present to the respondent. [13]

All of the pilot interviews used the attributes and levels described above. However, there were two additional treatments in the experimental design. First, the design included administering the questionnaire in French and English. Having these two versions allowed us to interview French-speaking Canadians. In addition, two versions varied the instructions given to respondents. In Version A, respondents were allowed to make their own assumptions about how to account for illness-related lost wages. In Version B, respondents were told to assume that "any missed time from work will be covered by paid sick leave." The purpose of these treatments was to test how respondents would react to the stated costs under alternative instructions. Specifically, this part of the design tests whether respondents recoded the cost information in the survey to fit their own circumstances when no sick leave assumption is specified, because the recoding makes interpreting the WTP estimates more difficult.

All of the questionnaire versions use a computerized format programmed in Visual Basic™. Each version of the computerized questionnaire has several sections. The first section, the introduction, is designed to introduce respondents to the general topic of health and prepare them for the rest of the survey. Section 2 asks respondents to read a two-page article on heart and lung illnesses and then complete four quiz

[13] The morbidity design was limited to 12 comparisons primarily because of time and attention constraints of the respondents. Given that we also wanted to do an experimental section of mortality paired comparisons and we needed to limit the interview to 30 minutes, 12 was the maximum number of morbidity comparisons possible. We did some simulated draws of 12 comparisons from our design to ensure the 12 comparisons would provide sufficient coverage of the attribute combinations in our design.

questions.[14] The next section of the survey asks respondents to rate their own health using several attributes from the SP exercise to familiarize respondents with the range of attributes that will be used in the SP exercise. Section 4 of the survey contains the morbidity SP exercise. This section also explains the payment vehicle and other key terms, presents an example of a rated pair, and introduces the information treatment for paid sick leave. Following the morbidity section was an experimental section on mortality valuation asking respondents to rate five mortality pairs.[15]

In addition to these sections, the survey also collects health-history information on the respondent's personal health history and the health history of family members. Finally, the survey contains sociodemographic questions about age, gender, education, employment, paid sick leave, income, and the number of adults and children in the individual's household. These questions are included to develop a profile of respondents to use in the analysis of respondents' individual SP ratings.

In developing the survey instrument and its design, pretesting was conducted on the English version of the survey, including two focus groups and two rounds of one-on-one pretesting. The survey was pretested using both symptomatics[16] and the nonsymptomatics. The English version of the survey was translated into French and consultants from CROP, Inc. and Cogesult, Inc. verified the translation and examined the survey for biases (Joubarne and Barbeau, 1996).

Using the English and French survey instruments, we conducted a pilot test in March 1996. The self-administered, computerized survey was approximately 30 minutes long, and the incentive payment for each respondent was $10. A total of 246 surveys

---

[14] The incorporation of quiz questions in the survey design responds, at least in part, to the NOAA Blue Ribbon Panel's recommendation that survey practitioners evaluate whether respondents comprehend the commodity being valued (see 58 *Fed. Reg.* 4613).

[15] The results of this experimental mortality section are not discussed in this paper. For a discussion of these results, see Desvousges et al. (1996).

[16] For this study, symptomatics were defined as any respondent who had ever been diagnosed with asthma, lung infections, chronic bronchitis, emphysema, or heart disease.

were completed during the pilot test in Toronto and St. Hubert, a suburb of Montreal. Thus, with each of the 246 respondents[17] providing 12 ratings of the paired comparisons, there was a reasonable sample size for estimating the model for the purposes of a pilot analysis.

PILOT ANALYSIS

The primary objective of the pilot survey was to evaluate how well various elements of the survey instrument and survey design performed for pilot-survey subsamples, not to estimate definitive WTP values. In addition, the pilot survey was designed to evaluate the effects on WTP values of specifying sick leave versus not specifying sick leave. This section examines results of the pilot test in light of these objectives.

In our design, the symptom attribute is held constant for both the left and right bundle for each rated pair. Therefore, the symptom must be interacted with another attribute that varies within each rated pair in order for the symptoms to be used in the estimation. For the models presented in this section, we interact symptom with episodes. The symptom coefficients indicate the effect of a symptom-episode combination on health-state utility, holding the remaining health attribute, activity limitation, constant. Similarly, the activity-limitation coefficients indicate the effect of an activity-limitation level on health-state utility, holding symptom and episode constant. To avoid the dummy-variable trap, we omit the activity level "With no limitations" (NOLIM).[18] (See Table 2 for the definitions of variables used in this analysis.) Thus each activity-limitation coefficient is the effect on utility of a given activity limitation relative to NOLIM.

---

[17] After removing respondents with insufficient or inconsistent data, the SP models presented are based on 223 respondents.

[18] Because the symptom dummy variables are interacted with episodes, no symptom dummy variable needs to be omitted.

As mentioned in the previous section, the experimental design precludes particular attribute combinations from appearing in the paired comparisons in order to make the bundles credible to respondents. For example, the relatively mild symptom "stuffy/runny nose and sore throat" (NOSE) never occurs in combination with the "in hospital" (INHOSP) activity limitation. Experimental design constraints also limit some symptom-episode combinations. For example, in our design the NOSE symptom occurs with three, four, or five episodes, while WEAK occurs with one, two, or three episodes. Our preliminary models do not explicitly account for daily activity and episode restrictions from the experimental design in the model specification, but implicitly assume that all combinations can occur.

Table 2.
Variables Used In Morbidity Marginal Utility of Money Models

VARIABLES	DESCRIPTION
NOSE	Stuffy or runny nose and sore throat
EYE	Eye irritation
WEAK	Generally tired and weak
FLUTTER	Fluttering in chest and feeling light-headed
COUGH	Coughing, wheezing, shortness of breath
ACHE	Coughing or wheezing with fever, chills or aching all over
SWELL	Shortness or breath, and swelling in ankles and feet
PAIN	Pain in chest or arm
EPISODES	Number of Episodes per year (1 to 5)
SOMELIM	Activity Level 2: You can go to work, go to school, do housework, and participate in social or recreational activities, but you have some physical limitations (trouble bending, stooping, or doing vigorous activities) because of this health condition.
SOCIALIM	Activity Level 3: You can go to work, go to school, do housework, but you have some physical limitations (trouble bending, stooping, or doing vigorous activities), and cannot participate in social or recreational activities because of this health condition.
ATHOME	Activity Level 4: You cannot leave your house, go to work, go to school, do housework, participate in social or recreational activities, and you have some physical limitations (trouble bending, stooping, or doing vigorous activities) because of this health condition, but you can care for yourself.
NEEDHELP	Activity Level 5: You cannot leave your house, go to work, go to school, do housework, participate in social or recreational activities, and you need help caring for yourself (feeding, bathing, dressing, toilet) because of this health condition.
INHOSP	Activity Level 6: You are in hospital and need help caring for yourself (feeding, bathing, dressing, toilet).
LNCOSTCONST	Log of the cost levels per year (Can$50–Can$700) as a constant
FRENCH	Dummy variable = 1 for the French version of the survey
SCORE	Quiz score (percent correct)
CONJTIME	SP exercise completion time
AGE	The midpoint of the age category
EDUCATION	Number of years of education
MALE	Dummy variable = 1 if respondent is male
INCOME	The midpoint of the income category
SYMPTOMATIC	Dummy variable = 1 if the respondent is symptomatic
LOWHEALTH	Dummy variable = 1 if respondent rated his/her own health as fair or poor
LOWINFO	Dummy variable = 1 if respondent acquires health information an hour per month or less
PAIDLEAVE	Dummy variable = 1 if respondent has paid sick leave
SICKVERSION	Dummy variable = 1 if respondent took paid sick-leave version of the survey
NEITHER	Dummy variable = 1 if respondent does not have paid sick leave and did not take the paid sick-leave version of the survey

While we have not incorporated the episode and activity-level constraints into the models shown in Table 3, we have accounted for a similar problem with cost levels that occur only with certain symptoms. Cost enters these models in a log form, rather than linear. Using the log of cost allows the relative utility of health and money to vary across the symptom-specific ranges of costs presented to respondents, resulting in symptom-specific marginal utility of money estimates.

Design restrictions must be kept in mind when interpreting model coefficients. For the five symptoms that occur in combination with the omitted activity level "No limitations" (i.e., NOSE, EYE, WEAK, FLUTTER and COUGH), the ordered-probit symptom-episode coefficients shown in Table 3 indicate the disutility of one additional episode of the symptom with no activity limitations.[19] For example, one additional episode of "Generally tired and weak" (WEAK) reduces utility by 0.1096 for the English-speaking subsample.

The disutility of one additional episode of a given symptom with more severe activity limitations is calculated by adding the given activity-limitation coefficient to the symptom coefficient. For the three symptoms that never occur with the NOLIM activity level (ACHE, SWELL, and PAIN), the symptom coefficient must be added to one of the activity-limitation coefficients. Because the mildest limitation in our design for ACHE is SOMELIM, the smallest possible decrease in utility can be calculated by summing the coefficients for ACHE and SOMELIM: -0.0187 + -0.2016 = -0.2203 (for the English-speaking subsample).

[19] Utility units are arbitrarily scaled and must be interpreted relative to one another.

Table 3.
Morbidity Marginal Utility of Money Models

Variable	ENGLISH VERSION Coefficient	P-value	FRENCH VERSION Coefficient	P-value	NO SICK LEAVE SPECIFIED VERSION Coefficient	P-value	PAID SICK-LEAVE VERSION Coefficient	P-value
NOSE*EPISODES	-0.1464 *	0.094	0.0855	0.404	0.0553	0.580	-0.1343	0.122
EYE*EPISODES	-0.0831	0.172	-0.0584	0.402	-0.0095	0.880	-0.0988	0.140
WEAK*EPISODES	-0.1096 **	0.030	-0.0694	0.180	-0.0990 **	0.048	-0.0640	0.222
FLUTTER*EPISODES	-0.0367	0.384	-0.0654	0.138	-0.1084	0.010	0.0101	0.816
COUGH*EPISODES	-0.1121 ***	0.004	-0.0130	0.764	-0.0483	0.260	-0.0851 **	0.030
ACHE*EPISODES	-0.0187	0.692	0.0207	0.718	-0.0213	0.682	0.0215	0.670
SWELL*EPISODES	-0.1003 *	0.056	-0.1047 **	0.038	-0.1155 **	0.020	-0.0941 *	0.068
PAIN*EPISODES	-0.1642 *	0.102	0.0507	0.606	-0.0708	0.422	-0.0212	0.838
SOMELIM	-0.2016 ***	0.004	-0.2775 ***	0.000	-0.2733 ***	0.000	-0.1770 ***	0.014
SOCIALIM	-0.3710 ***	0.000	-0.3708 ***	0.000	-0.4903 ***	0.000	-0.2273 ***	0.000
ATHOME	-0.5547 ***	0.000	-0.2813 ***	0.000	-0.5229 ***	0.000	-0.2864 ***	0.000
NEEDHELP	-0.6268 ***	0.000	-0.4630 ***	0.000	-0.6117 ***	0.000	-0.4350 ***	0.000
INHOSP	-0.7406 ***	0.000	-0.5209 ***	0.000	-0.7375 ***	0.000	-0.4914 ***	0.000
LNCOSTCONST	-20.8477	0.330	-20.8857	0.346	16.4925	0.474	-33.2674 *	0.070
FRENCH					-0.3414	0.950	4.7424	0.384
SCORE	-0.2560 ***	0.006	-0.2780 **	0.030	-0.2118 **	0.050	-0.3383 ***	0.000
CONJTIME	-0.4076	0.302	-0.4161	0.490	-0.4541	0.476	-0.2784	0.494
AGE	33.6044 *	0.072	37.9891 *	0.070	26.3357	0.198	36.9087 *	0.056
EDUCATION	0.4927	0.692	-0.2842	0.828	-1.0679	0.406	1.3566	0.250
MALE	-0.9223	0.852	-5.1248	0.366	0.3559	0.950	-6.2055	0.218
INCOME	-0.0800	0.362	0.0001	0.998	-0.0206	0.862	0.0178	0.848
SYMPTOMATIC	-9.6873 *	0.064	8.6808	0.138	-2.3135	0.676	0.1557	0.976
LOWHEALTH	-4.1628	0.598	-9.8663	0.120	-13.1965 *	0.062	0.5777	0.936
LOWINFO	-3.9513	0.190	3.7046	0.242	-5.7260 **	0.048	2.8350	0.330
PAIDLEAVE	9.4427	0.306	9.4654	0.384	-7.8395	0.236	2.9092	0.680
SICKVERSION	2.2947	0.754	4.7077	0.614				
NEITHER	16.9340	0.156	16.0655	0.272				
ALPHA1	-1.2373	0.000	-1.4210	0.000	-1.3247	0.000	-1.3437	0.000
ALPHA2	-0.2607	0.000	-0.2037	0.000	-0.2506	0.000	-0.2390	0.000
ALPHA3	0.6672	0.000	0.4796	0.000	0.5673	0.000	0.5517	0.000
Like. Ratio Chi-sq.	180.583		94.279		142.906		119.163	
Prob(Chi-sq.)	0.000		0.000		0.000		0.000	
Maddala pseudo R2	0.119		0.070		0.100		0.083	

Neglecting the constraints imposed by the experimental design introduces some degree of bias in the symptom and activity coefficients. This bias may be particularly pronounced for the PAIN symptom, which occurs with one to three episodes and the three most severe activity limitations. In contrast, NOSE and EYE never occur with the two most severe activity limitations. These interaction patterns suggest that treating symptom and activity effects as additions is inappropriate, and biases utility estimates to some degree. The complicated pattern of restrictions makes it difficult to predict the direction or magnitude of the biases.

More complicated specifications that allow for nonlinearities and more complicated interactions are possible. Such refinements have not been explored at this stage of the study. Instead, we have used linear specification for the symptoms, episodes, and activity levels to test whether pilot-survey respondents in various subsamples detected significant differences among health states and whether signs of the estimated effects are consistent with theory and logic. These models are sufficient to diagnose any potentially serious problems in our experimental design or survey questionnaire, which is the goal of this pilot study.

Symptoms were chosen on the basis of policy relevance, not necessarily perceived salience on the part of respondents. Nevertheless, respondents in the English-speaking subsample perceived statistically significant utility losses for:

- Stuffy or runny nose and sore throat (NOSE)
- Generally tired or weak (WEAK)
- Coughing, wheezing, shortness of breath (COUGH)
- Shortness of breath and swelling in ankles and feet (SWELL)
- Pain chest or arm (PAIN)

All symptom coefficients have the expected negative sign, indicating each symptom decreases utility. The French-speaking subsample was much less sensitive to symptoms, with only SWELL being statistically significant and three coefficients having the wrong sign. With English-speaking and French-speaking respondents assigned randomly to sick-leave version treatments, there is no clear effect of version on symptom salience.

Thus for most symptoms, the English-speaking subsample results support the feasibility of estimating meaningful WTP values for policy-relevant respiratory and cardiac symptoms in a full-scale study. For example, the coefficient indicates a utility loss of 0.1096 for an episode of "Feeling generally tired and weak" with no limitations on daily activity. This utility loss can be rescaled to dollars using the estimated marginal utility of money for WEAK of 0.0843 utility units per \$100.[20] Thus, a health-state change of 0.1096 utility units corresponds to a WTP of $\$100 \cdot \frac{0.1096}{0.0843} = \$130$.

All significant symptom coefficients have negative signs across all four subsamples, indicating these symptoms result in a loss of utility. There is no natural ordering of symptom disutility, so we have no general expectations about relative magnitudes. Nevertheless, we might expect a larger difference between NOSE (-0.146 in the English-speaking subsample) and PAIN (-0.164). Recall, however, that the mildest activity limitation for NOSE is no limitations, while the mildest activity limitation for PAIN is confined at home but able to care for self (ATHOME). Thus the two coefficients are not directly comparable because this preliminary model specification does not account for symptom-activity restrictions.

The coefficients for the English-speaking subsample for the five symptoms which allow the NOLIM activity level, range between -0.04 for FLUTTER and -0.15 for NOSE. These differences are statistically insignificant, as are the corresponding point estimates for the French-speaking subsample. The statistical similarity of these

[20] We evaluated δ^i at the means of the explanatory variables and cost at the midpoint of the range of cost levels for each symptom.

coefficients indicates that we cannot reject the hypothesis that these symptoms provide an equivalent loss in utility for the pilot-survey sample sizes.

Dividing the sample according to the sick-leave version shows a similar pattern. For the subsample where paid sick leave was specified, two of the seven symptom coefficients are significant at conventional levels and two additional symptoms are close to significant. Three of the symptom coefficients are significant for the subsample where no sick-leave information was specified. Again, large standard errors make differences among the coefficients in both subsamples statistically insignificant.

The relative insensitivity of pairwise ratings to symptom differences has several possible explanations. First, there is no reason to suppose that respondents hold strong preferences among the policy-relevant symptoms included in this study apart from the effect that symptoms have on daily activities. Thus we would not expect to see highly significant differences among relatively mild symptoms, holding daily activity level and number of episodes constant. Alternatively, small subsample sizes of about 110 observations may have affected our ability to detect significant differences.

Finally, it is possible that some of the observed insensitivity was induced by holding symptom constant for each pairwise comparison. While this strategy was adopted to reduce respondents' cognitive burden and to simplify the experimental design, it may have caused them to focus on other attributes of the pairs. We believe it is worth investigating the possibility of varying symptoms within each rated pair for the full-survey implementation. While this change would increase the complexity of the study design, it also would increase modeling flexibility. It also would reduce any insensitivity to symptom induced by the design itself.

The results for the daily activity levels are consistently strong across all four subsamples. The strength and regularity of these results clearly indicate that respondents evaluated activity limitations in a logical way. Either respondents found it easier to rate the activity levels than the symptom levels, or this attribute was more comprehensible and/or important. In either case, these results indicate the experimental design was successful in obtaining meaningful stated preferences for these activity levels.

All activity-limitation coefficients are negative and significant at the highest level for all subsamples. Moreover, for all subsamples except French-speaking respondents, the magnitude of the utility differences between the excluded daily activity level (no limitations) and included activity levels increases monotonically with more severe limitations, as expected. For example, the English-speaking estimate of the utility loss associated with limitations on social and recreational activity (SOCIALIM) is -0.371. The utility loss for being confined to home (ATHOME) is -0.555. Thus being confined to home results in a larger utility loss than facing only social and recreational limitations. This difference is statistically significant and corresponds to a difference in WTP of about $220 for one episode of the WEAK symptom.

The only exception to the direct relationship between disutility and limitation severity occurs in the estimates for the French-speaking subsample. The coefficients for SOCIALIM and ATHOME for the French subsample are -0.371 and -0.281, respectively, suggesting that ATHOME is a less severe restriction than SOCIALIM.[21] Nevertheless, these coefficients are not significantly different from each other. The lack of sensitivity to the severity of the activity limitation in the French-speaking sample may be a matter of some concern, especially given the clearly consistent patterns in the other three subsamples.

All of the activity-level coefficients for the subsample, given no assumption about sick leave, are larger than the corresponding coefficients for the paid sick leave

[21] Note that the English and French estimates for SOCIALIM appear virtually identical, -0.3710, and -0.3708, respectively. It is inappropriate, however, to compare absolute marginal utility estimates across samples. The scale for the subsamples is different, as indicated by the difference in the marginal utility of money estimates, 0.084 and 0.056 for WEAK. The same attribute coefficient corresponds to different WTP values of $440 and $660.

subsample. The only difference between the two subsamples that is not statistically significant is for SOMELIM. Thus respondents in the subsample without the paid sick leave assumption appear to have expressed larger losses than corresponding respondents who were told to assume they had paid sick leave. This result indicates that the respondents who were left to make their own assumptions about lost wages adjusted the costs presented in the rated pairs upward by some unknown amount. These respondents, therefore, did not interpret the cost levels as the total price difference for the SP health-state differences.

The marginal utility of money was modeled as a linear function of 13 variables plus a constant term. Table 4 shows the mean or median values for these characteristics across the four subsamples. As expected, there are no significant differences across the two sick-leave versions. These two versions were assigned randomly, so we should see no differences across these subsamples.

Comparing across the two language subsamples, however, we see several significant differences. Specifically, the French-speaking respondents have:

- Longer SP exercise times
- Lower median quiz scores
- Lower median education levels
- Lower median income levels
- Less coverage by paid sick leave

Because each subsample is modeled separately, any differences in the effect of these variables on the marginal utility of money will be reflected in the model coefficients. In addition, the calculation of the marginal utility of money is conditional on the means of these variables, so these differences will be reflected in the mean marginal utility of money for each subsample.

Table 4.
Characteristics of Respondents Used In The Model: Screened Sample, By Language, and By Sick-Leave Version

	LANGUAGE VERSION		SICK-LEAVE VERSION	
Characteristic	**English**	**French**	**Version A** (No Information)	**Version B** (Paid sick leave)
Language version (percent French)	0%	100%	50%	46%
Median quiz score	75%	50%	75%	50%
	Significantly different at the 1 percent level			
Mean SP exercise completion time	6.1 minutes	7.5 minutes	6.6 minutes	7.0 minutes
	Significantly different at the 1 percent level			
Median age (category in years)	40 – 49	40 – 49	30 – 39	40 – 49
Median education (category)	Completed community college, technical college, CEGEP, or nursing program	Completed secondary or high school	Some community college, technical college, CEGEP, or nursing program	Some community college, technical college, CEGEP, or nursing program
	Significantly different at the 1-percent level			
Sex (percent male)	50%	49%	52%	46%
Median income (category)	\$30,000 – \$39,000	\$20,000 – \$29,000	\$30,000 – \$39,000	\$30,000 – \$39,000
	Significantly different at the 5-percent level			
Symptomatic[a] status (percent symptomatic)	37%	39%	38%	37%
Percent who rated their health fair or poor	9%	15%	12%	11%
Percent who seek health information an hour per month or less	62%	53%	62%	54%
Percent who have paid sick leave	38%	21%	26%	34%
	Significantly different at the 10-percent level			
Number of observations	119	109	113	115

Few of the personal characteristic variables are significant for any of the models.[22] A negative coefficient means that an increase in the variable results in an increase in the disutility of cost compared to nonsymptomatics. The negative coefficient, therefore, indicates that an increase in the variable increases the utility of money relative to the utility of health, and thus results in lower WTP estimates.

Symptomatic respondents (SYMPTOMATIC) have increased disutility of costs compared to nonsymptomatics for the English-speaking subsample. This result indicates that symptomatic respondents have lower WTP for health states. If the symptomatic and nonsymptomatic groups have the same utility function for health, as we have modeled them, this result violates the concept of diminishing marginal utility for health. Symptomatic people presumably have a lower baseline health state and thus should be willing to pay more for a marginal change in health than nonsymptomatic people at higher levels of baseline health. This result, therefore, may indicate the need to model these two groups separately, allowing them to have different-shaped utility functions. The sample size of symptomatics in the pilot study is too small to allow for this type of analysis. Nevertheless, the design of the full study, as discussed in Desvousges et al. (1996), would allow for such investigations.

Increases in age tend to increase WTP across all subsamples, although the coefficient is not significant for the subsample receiving the version with no assumption about sick leave. Higher WTP on the part of older respondents is consistent with a greater interest in health among this group. The positive sign on the NEITHER dummy variable for both the English-speaking and French-speaking subsamples is consistent with respondents in this group perceiving that their costs would be higher than those specified in the survey. The significant negative coefficients for the respondent's quiz score

[22] For models in which the marginal utility of money is specified as a single coefficient on COST (not reported here), the parameter is statistically significant with correct sign for all models and subsamples. This result confirms that respondents successfully traded off the cost attribute against other attributes and justifies using the cost coefficient as the marginal utility of money in WTP calculations.

(SCORE) across all four subsamples indicate that increases in quiz score increase the disutility of costs and reduce WTP. We had no particular expectations about the sign for the SCORE variables.

The coefficient on the dummy variable indicating that the respondent has paid sick leave (PAIDLEAVE) is negative for the NO-ASSUME subsamples, and positive for the ASSUME subsample, but significant in neither. The negative sign indicates that respondents who do not have paid leave have a smaller marginal utility of money and thus lower WTP than those who do, other things held constant. However, a more plausible explanation is that respondents without paid leave tended to adjust the cost levels shown in the SP profiles to incorporate additional costs from lost wages. This recoding dilutes the sensitivity to specified cost levels for these respondents, making it appear as if they get less utility from an additional dollar. Thus, recoding the cost levels undermines the validity of the marginal utility of money estimate based on this coefficient. In effect, the PAIDLEAVE coefficient for the subsample who received no instruction on how to treat unpaid sick leave is the marginal effect of a dollar of specified costs plus some unobserved individual-specific adjustment.

We noted earlier that the NO-ASSUME subsample expressed significantly larger utility losses for activity limitations than the ASSUME subsample. In addition, the marginal utility of money for the NO-ASSUME subsample is smaller across symptoms. For example, the mean marginal utility of money is 0.060 per $100 for WEAK, compared to 0.077 for the ASSUME subsample. The relatively smaller marginal utility of money values serves to magnify the monetary-equivalent utilities for activity-limitation losses.

This result confirms that respondents seriously considered cost factors in evaluating relative utilities of health states. It also is remarkable that inclusion of a single sentence on paid sick leave has such significant effects on stated preferences. While the availability of paid sick leave should affect respondents' ratings of alternative health states, it is reassuring that respondents paid attention to this important detail.

Nevertheless, the significant differences between the treatment subsamples suggest that not standardizing the influence of having or not having paid sick leave may compound the effects of monetary and nonmonetary factors on stated preferences. Thus, the final survey should clearly state what the role of paid sick leave is and collect information on respondents' actual sick-leave benefits.

As discussed above, the models presented in this paper are designed to diagnose problems in the survey instrument prior to the full administration. These models are not designed to give definitive estimates of willingness to pay. Nevertheless, these models do result in WTP estimates that are illustrative. These estimates, however, are preliminary and should be interpreted cautiously. They are based on relatively small samples and are calculated from simplified models that do not incorporate experimental-design restrictions and other important refinements that may affect the estimates.

Figure 2 shows the mean WTP and corresponding 90-percent confidence intervals for one episode of the shortness of breath with swelling in ankles and feet symptom, for each of the activity levels used for that morbidity symptom. The "no limitations" activity level never was paired with this symptom, so the values shown begin with the second activity level category of some physical limitations. In Figure 2, we compare the WTP estimates for the English-speaking and French-speaking subsamples.

The confidence intervals for the English-speaking respondents are about half as wide as those for the French-speaking respondents.[23] Also, as discussed above, the activity-level coefficients decrease systematically for the English-speaking subsample as the limitations become more serious, resulting in WTP estimates that increase systematically across the limitation categories. However, for the French-speaking subsample, the coefficient for ATHOME is larger than for SOCIALIM, so the WTP estimate for ATHOME is less than the estimate for SOCIALIM ($790 versus $970).

[23] Confidence intervals were bootstrapped from distributions defined by the parameter and variance-covariance estimates using 3,000 repetitions.

With the large confidence intervals for the French-speaking respondents, the difference between these two values is not statistically significant.

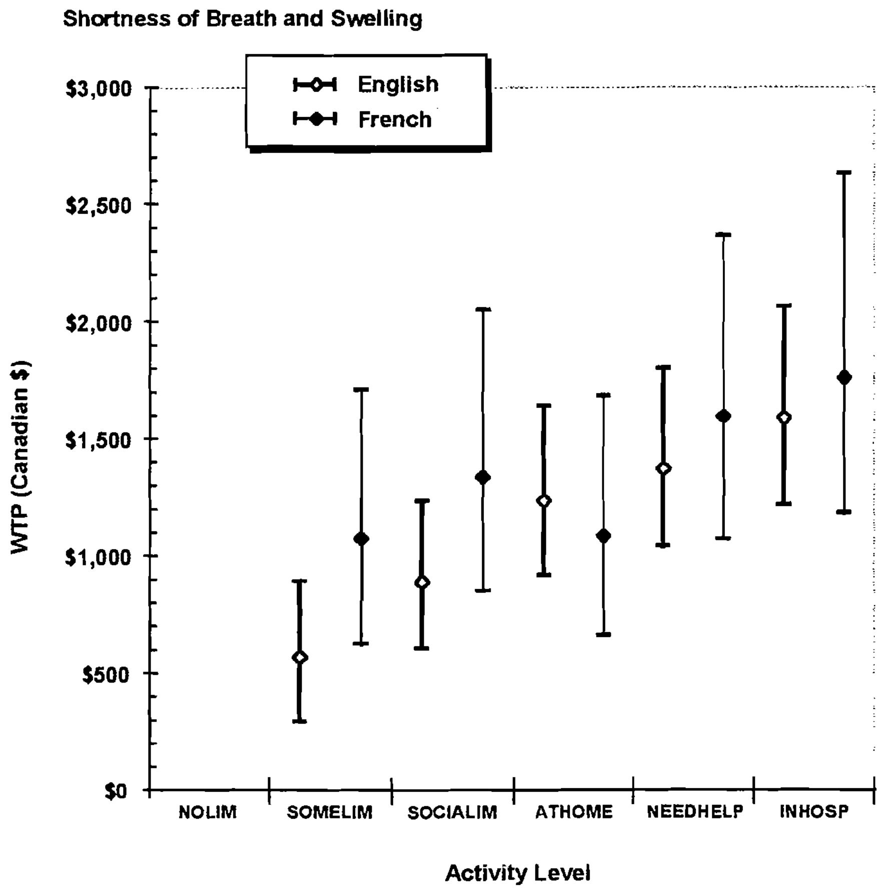


Figure 2.
WTP Confidence Intervals:
Shortness Of Breath And Swelling

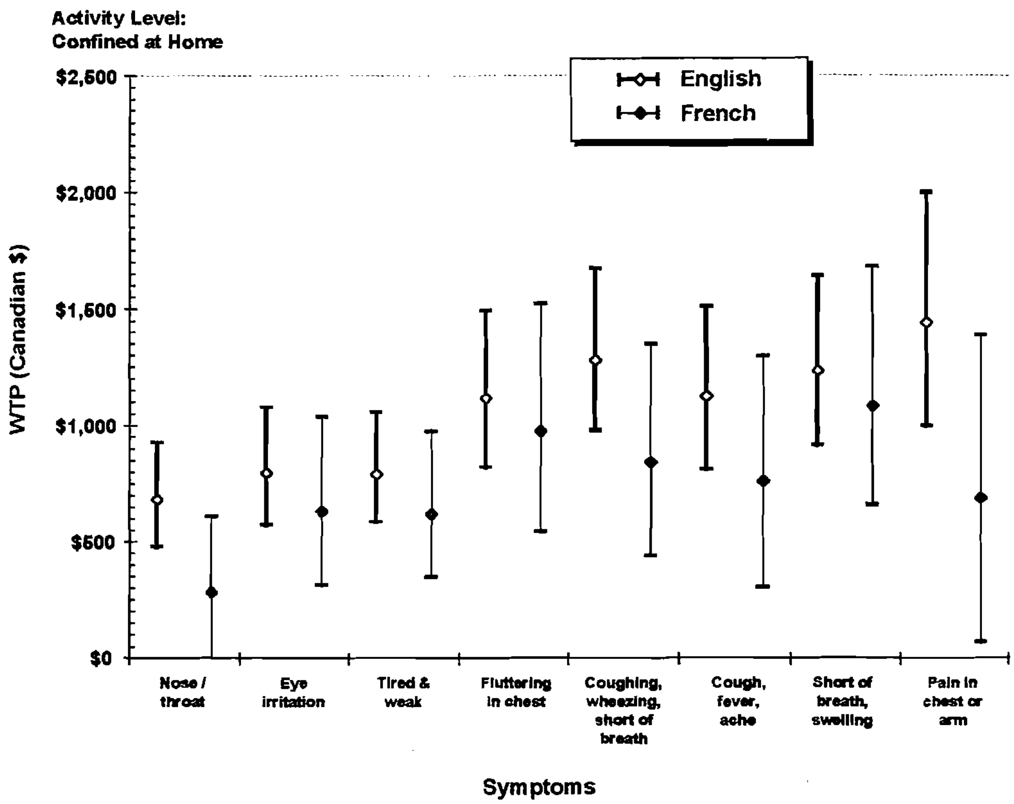
Comparing the English-speaking estimates across activity levels, adjacent estimates are in general not significantly different, except for SOCIALIM limitations compared with ATHOME. However, when comparing activity levels more than one level apart, the differences almost always are statistically significant.[24] In contrast, the

[24] This result may simply be an artifact of our small sample sizes.

only comparison of daily activity levels that yields a statistically significant difference for the French-speaking respondents is the second level (SOMELIM) and INHOSP.

In Figure 3, we present the WTP estimates from a different perspective. In this figure, we show the WTP point estimates and 90-percent confidence intervals for all. eight symptoms for one activity level (confined at home), comparing the English-speaking and French-speaking subsamples. Once again, the confidence intervals for the French-speaking estimates are roughly twice as wide as the English counterparts. The other important result from this figure is that, while there is some variation in the means across these symptoms, the symptom that is significantly different within a subsample is Nose/throat for the French subsample. For example, for the English-speaking subsample, the symptom with the highest mean is "pain in chest or arm" ($990), while the smallest mean occurs for "cough, fever and ache" ($790). This difference is not statistically significant. Thus, as indicated by the model results above, there is not much variation across symptoms.

Figure 3.
WTP Confidence Intervals for All
Eight Symptoms: Confined At Home



It is often useful to compare WTP estimates with values estimated in studies of similar commodities. However, the health-valuation literature consists largely of older contingent-valuation studies, many with serious problems. (See Desvousges et al., 1996, for a review of this literature.) In Johnson, Fries, and Banzhaf (1996), these studies are combined into a meta-analysis. This meta-analysis uses the QWB health-status index as a mechanism for combining information from dissimilar studies. In essence, the health states valued in each study are converted to QWB ratings. These QWB ratings then are used, along with other important information from the models, as independent variables in a WTP regression model. The resulting coefficient estimates can be used to predict WTP for any health state that conforms to the QWB classification system.

To compare our pilot study WTP estimates to this literature, we used the meta-analysis equation from Johnson, Fries, and Banzhaf (1996) to predict WTP for each of our eight symptoms. To reflect the range of daily activity levels, we calculated estimates for the mildest activity limitation and the most serious activity limitation associated with that symptom in our design.

Table 5 compares these meta-analysis estimates with estimates for the same symptom/activity level combinations from our pilot study data. Table 5 presents only estimates from the English-speaking subsample for comparison purposes.

The comparisons shown on Table 5 reveal some interesting patterns. First, for the more severe symptoms (COUGH, ACHE, SWELL, and PAIN) with the most severe activity limitations, the estimates from our pilot data are very similar to the estimates from the QWB meta-analysis. However, for the mild activity level limitations, in all cases except fluttering in chest, the pilot-test WTP estimates are substantially higher than the meta-analysis estimates. Also, for the mild symptoms coupled with the severe activity limitations, the pilot test estimates far exceed the meta-analysis results. The meta-analysis values are by no means a criterion standard with which our results must be consistent. Nevertheless, the comparison suggests that our preliminary estimates may overstate WTP, especially for mild symptoms and modest activity restrictions. Additional utility-theoretic and study-design features need to be incorporated into these models to obtain more reliable WTP estimates. Furthermore, small sample sizes and convenience sampling make it inappropriate to use these estimates for any purpose other than as diagnostics in evaluating the general performance of the survey instrument.

Table 5.
Comparison Of WTP Estimates With QWB Meta-Analysis Estimates

Symptom	CANADA PILOT STUDY		QWB META-ANALYSIS	
	Mild Activity Limitation	Severe Activity Limitation	Mild Activity Limitation	Severe Activity Limitation
Stuffy/runny nose and sore throat	$143	$683	$38	$203
Eye irritation	$104	$797	$65	$351
Generally tired and weak	$130	$875	$85	$694
Fluttering in chest and feeling light-headed	$69	$1,471	$123	$1,513
Coughing, wheezing, and shortness of breath	$215	$1,638	$84	$1,029
Coughing or wheezing with fever, chills, or aching all over	$433	$1,491	$146	$1,029
Shortness of breath, swelling in ankles and feet	$570	$1,586	$215	$1,513
Pain in chest or arm	$1,443	$1,816	$663	$1,513

CONCLUSIONS

We conclude from our statistical analysis of the pilot-test data that most of the survey-design features resulted in sensible patterns of responses with a few notable exceptions. Respondents generally were appropriately attentive to differences in attribute levels, accepted trade-offs between cost and health at the cost levels specified, and accounted for differences in information about paid sick leave in expected ways. There is evidence that the version with no information about paid sick leave led to confounding of health and income-loss effects. This reaction complicates deriving and interpreting WTP estimates. Thus we recommend use of the specified sick-leave version of the survey for the full implementation.

There is substantial empirical evidence that the English-speaking subsample was more sensitive to symptom differences and provided ratings that were less noisy than the French-speaking subsample. Coefficient estimates from the French-speaking subsample were more frequently insignificant and anomalous than the English-speaking estimates.

Nevertheless, the majority of French-speaking respondents provided coherent, usable ratings. In a separate analysis, not reported here, we found 20 to 25 percent of this group's ratings were unacceptably noisy. While we have no satisfactory explanation for the noisier responses among the French-speaking subsample, we have developed procedures for identifying problem observations. These observations can be deleted, which may improve the precision of coefficient estimates and narrow WTP confidence intervals.

REFERENCES

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

Cattin, Philippe, and Dick R. Wittink. 1982. "Commercial Use of Conjoint Analysis: A Survey." *Journal of Marketing* 46:44–53.

Desvousges, William H., F. Reed Johnson, Sara P. Hudson, Alicia R. Gable, and Melissa C. Ruby. 1996. *Using Conjoint Analysis and Health-State Classifications to Estimate the Value of Health Effects of Air Pollution: Pilot Test Results and Implications.* Draft report prepared for Environment Canada, Ontario Hydro, Ontario Ministry of Environment and Energy, and Quebec Ministry of Environment. Durham, NC: Triangle Economic Research.

Gan, Christopher, and E. Jane Luzar. 1993. "A Conjoint Analysis of Waterfowl Hunting in Louisiana." *Journal of Agricultural and Applied Economics* 25:36–45.

Johnson, F. Reed, Erin E. Fries, and H. Spencer Banzhaf. 1996. *Valuing Morbidity: An Integration of the Willingness-to-Pay and Health-Status Index Literatures.* TER Technical Working Paper T-9601. Durham, NC: Triangle Economic Research.

Johnson, R. Reed, William H. Desvousges, Erin E. Fries, and Lisa L. Wood. 1995. *Conjoint Analysis of Individual and Aggregate Environmental Preferences.* TER Technical Working Paper T-9502. Durham, NC: Triangle Economic Research.

Joubarne, Luc, and Louis-Philippe Barbeau. 1996. *Feasibility of the Administration in Quebec of a Questionnaire Translated From English Into French.* Final Report. Quebec: Cogesult Inc. and CROP Inc.

Keeney, Ralph L., and Howard Raiffa. 1978. Decisions With Multiple Objectives: Preferences and Tradeoffs. New York: John Wiley and Sons.

Krupnick, Alan J., and Raymond J. Kopp. 1988. *The Health and Agricultural Benefits of Reductions in Ambient Ozone in the United States.* Discussion Paper QE88-10. Washington, DC: Resources for the Future.

Mackenzie, John. 1993. "A Comparison of Contingent Preference Models." *American Journal of Agricultural Economics* 75(3):593–603.

National Oceanic and Atmospheric Administration. 1993. "Proposed Rules: Panel Report." *Federal Register* 58(10):4601–4614.

Opaluch, James J., Stephen K. Swallow, Thomas Weaver, Christopher W. Wessells, and Dennis Wichelns. 1993. "Evaluating Impacts From Noxious Facilities: Including Public Preferences in Current Siting Mechanisms." *Journal of Environmental Economics and Management* 24:41–59.

Roe, Brian, Kevin J. Boyle, and Mario F. Teisl. 1994. "Deriving Compensating Variation from Conjoint Data Sets: A Theoretical Model and Empirical Issues." Draft Working Paper.

Train, Kenneth. 1995. *Simulation Methods for Probit and Related Models Based on Convenient Error Partitioning.* Working Paper 95-237, Department of Economics, University of California, Berkeley.

Viscusi, W. Kip, Wesley A. Magat, and Joel Huber. 1991. "Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Trade-offs for Chronic Bronchitis." *Journal of Environmental Economics and Management* 21:32–51.

Wittink, Dick R., and Philippe Cattin. 1989. "Commercial Use of Conjoint Analysis: An Update." *Journal of Marketing* 53:91–96.

Modeling Participation Activities that Require Experience: An Application to Whitewater River Recreation

Richard C. Ready and Donald Kemlage

Richard Ready is a Visiting Associate Professor, Department of Economics and Social Sciences, Agricultural University of Norway. Donald Kemlage is a GIS specialist, ESRI. This research was performed while both authors were at the Department of Agricultural Economics, University of Kentucky, and was supported by the Kentucky Agricultural Experiment Station. Data was provided under agreement with the Gauley River National Recreation Area.

Modeling Participation in Recreation Activities that Require Prior Experience: An Application to Whitewater River Recreation

Abstract

A zonal travel cost model of demand for whitewater recreation on the Gauley River, WV, is estimated both for private paddlers and for clients of commercial rafting companies. For private paddlers, who must have some prior experience before paddling this river, individuals with close access to many whitewater river sites are more likely to have that required experience, and therefore more likely to visit the Gauley River. For commercial rafters, who do not need prior experience, individuals with close access to many river sites are less likely to visit the Gauley River, due to substitution across sites. Failure to account for prior experience and learned skills will bias consumer surplus estimates for activities that require them.

Keywords: whitewater, recreation, travel cost model, learning

Modeling Participation in Recreation Activities that Require Prior Experience: An Application to Whitewater River Recreation

For many outdoor recreation activities, participation at a given site can only be done by individuals who have learned specific skills, have sufficient prior experience in the activity, and own appropriate equipment. Without those skills, experience and equipment, the activity would be unenjoyable and possibly dangerous. Examples include surfing, rock and mountain climbing, hang gliding, and whitewater paddling. The process of acquiring the experience, skill and equipment necessary to participate in these activities is commonly referred to as "taking up" the sport. An important part of "taking up" such sports is time spent learning activity-specific skills at suitable "beginner" sites. Special schools have emerged at such sites to help hopeful recreators gain the necessary skills. Taking up a sport such as mountain climbing or hang gliding can involve a quite substantial commitment of time and money before the participant is ready to enjoy the sport on his or her own.

For a given site that requires prior experience, the population of potential visitors qualified to visit the site is therefore self-selected. Only those individuals who have previously made a decision to take up the activity can visit the site. When modeling visitation to such a site, it is therefore important to understand the factors that influence that prior decision. Of particular importance for travel cost modeling is the role that is played by other recreation sites suitable for the activity. It has long been recognized that another recreation site can act as a substitute to a study site, so that visitation to the study site will be lower from areas close to the substitute site (Caulkins, Bishop and Bouwes; Rosenthal). When considering activities that require prior commitment to the activity and specialized skills, however, proximity to other sites will likely increase the probability that the activity will be taken up. This inducement to take up the activity will work in the opposite direction to the substitute effect, so that proximity to other similar sites could actually increase the probability of visiting a given site. As with substitute sites, failure to account for the role played by other sites in influencing whether the individual takes up the activity will bias estimated parameters in the visitation model, though the direction of the bias can be opposite to that associated with substitute sites.

As an example, consider mountain climbing. An individual who lives close to sites suitable for mountain climbing is more likely to take up the sport. So, for example, we expect to see more qualified mountain climbers per capita in Colorado than in Kansas. If we were to intercept visitors to a high profile climbing site such as Mount Rainier, in Washington State, we might expect to see more visits per capita from Colorado than from Kansas, even though Colorado has more sites that are close substitutes to Mount Rainier as a climbing destination. A zonal model of visitation that did not account for the influence of required skills and experience would attribute some or all of the difference in visitation rate to the difference in the cost of reaching Mt. Rainier from these two states, leading to a biased estimate of the slope of the visitation function, and biased estimates of the consumer surplus generated by Mt Rainier.

This study estimates demand models for whitewater paddling, an activity that requires specialized skills and equipment. The particular site for which demand models were estimated is the Gauley River, in central West Virginia. The Gauley River is considered suitable for advanced and expert paddlers only (Barrow). Thus, the population of potential visitors is limited to those individuals who have previously invested the time and money required to become an advanced paddler. An index of the availability of whitewater paddling opportunities is constructed and a reduced form model is estimated that accounts for the relationship between access to whitewater opportunities and the probability of taking up the sport. It is shown that because access to whitewater rivers increases the probability of taking up the sport, individuals who live near whitewater rivers are more likely to visit the Gauley River.

An identical model is estimated for clients of commercial rafting companies, who share this resource with private paddlers but who do not need prior experience. Here, other whitewater rivers can serve as substitutes to the Gauley River, so that individuals who live near whitewater rivers are less likely to visit the Gauley River.

Apart from insights gained from examination of the role of other sites on the probability of taking up the sport, estimates of the consumer surplus from whitewater paddling trips are useful in their own right for several ongoing policy issues. The Federal Energy Regulatory Commission (FERC) is currently in the process of relicensing existing hydropower dams, and is required by law to consider the impacts of dam operation on recreation. Estimates of consumer surplus from whitewater recreation would also be useful to dam management agencies such as the Bureau of Reclamation and the U.S. Army Corps of Engineers, when making trade-offs among power generation, lake recreation, irrigation, shipping, and in-stream flows for fishing and whitewater recreation, and when considering construction of new dams that would impact whitewater resources. More locally, estimates of consumer surplus for this particular river will be useful to the National Park Service, which manages access to the river, and the Army Corps of Engineers, which manages water flows in the river.

A Model of Site Visitation with Prior Learning

This study models visitation to a 24 mile section of the Gauley River from Summersville Dam to Swiss, WV. This section is one of the premier whitewater rivers in the eastern U.S. (Burrell and Davidson). It is paddled both by private paddlers, who own their own equipment (kayaks, canoes, and rafts) and guide themselves, and by commercial rafts, which typically hold eight to ten paying customers and one professional guide. On the International Scale of Whitewater Difficulty, which rates rivers on a six-class scale, this section is rated as class IV (advanced) and class V (expert) (Barrow). For a private paddler, it typically takes one to three seasons of experience before the paddler is skilled enough to paddle this section safely. These paddlers gain that experience on other, easier rivers. In contrast, clients of commercial rafting companies can paddle the Gauley River with no prior experience.

Visitation by private paddlers to a site like the Gauley River that requires prior experience should be modeled as a two-stage decision problem. First, the inexperienced potential user must decide to invest the time and money needed to learn to paddle. Those individuals who choose to learn those skills can then also choose to visit the Gauley River. The decision whether to take up the sport, purchase the required equipment, and learn the required skills depends in part on the cost of accessing suitable sites. Let C_W represents the cost of accessing whitewater rivers. This would include both beginner sites suitable for learning necessary skills, and more advanced sites where those skills can be used. The specific form of C_W will be considered later. There are other costs that will influence the decision whether to take up the sport, such as the cost of purchasing required equipment, but as these are not likely to vary systematically across individuals, they will not be explicitly modeled here.

An individual i who takes up whitewater paddling receives utility $U^P(M-C_W, S_i, \varepsilon_i^P)$, where M is wealth, S_i represents socio-demographic information about individual i. and ε_i^P is an error term that includes both random components and unobservable information about individual i. Components of ε_i^P could include special aptitude for the activity or personal information such as whether the individual's parents or close friends participate in the activity. If individual i does not take up whitewater paddling, then he or she receives utility $U^{NP}(M, S_i, \varepsilon_i^{NP})$. An individual chooses to take up the sport if

$$(1) \qquad U^P(M-C_W, S_i, \varepsilon_i^P) > U^{NP}(M, S_i, \varepsilon_i^{NP}).$$

Both sides of (1) contain random components, so it is impossible to completely determine who will have taken up the sport and who will not. However, for each individual there is some probability that he or she has taken up the sport, given by $p(C_W, S_i)$. An increase in C_W will decrease the probability that (1) is satisfied, so $\partial p/\partial C_W < 0$. Consequently, individuals who live close to whitewater rivers are more likely to be whitewater paddlers.

After an individual becomes a whitewater paddler, and gains sufficient skill, he or she can visit an advanced site like the Gauley River. On any given recreation occasion (for example a weekend), a paddler decides whether to visit the Gauley River based on

$$(2) \qquad u_i^G(m-C_G, S_i, \eta_i^G) > u_i^0(m, C_W, Si, \eta_i^0)$$

where u_i^G ( ) is the utility the individual receives if he or she visits the Gauley River on that occasion, u_i^0 ( ) is the utility the individual receives if he or she chooses some other activity on that occasion. m is the relevant budget constraint for that recreation occasion, C_G is cost of a visit to the Gauley River, and η_i^G and η_i^0 are error terms specific to that recreation occasion. Note that we distinguish between the long-term anticipated utility from becoming a paddler, given by $U^P(\ )$, and the single-occasion utility from a particular visit, given by u^G. C_W enters the single-occasion utility function for non-visitors because those non-visitors may choose to visit some other river on that recreation occasion. The conditional probability that individual i visits the Gauley River on a particular occasion, conditional on individual i having already taken up the sport, is given by $g(C_G, C_W, S_i)$. Higher C_G decreases the left hand side of (2), so that $\partial g/\partial C_G < 0$. Higher C_W decreases the right hand side of (2), so that $\partial g/\partial C_W > 0$. This latter result is the substitution effect that is widely recognized in the literature.

However, that substitution effect is moderated when considering the unconditional probability of visitation. Absent knowledge of whether individual i has taken up the sport, the probability that he or she will visit the Gauley River on a given recreation occasion is

(3) $$v(C_G, C_W, S_i) = g(C_G, C_W, S_i)\, p(C_W, S_i)$$

The first order partial derivatives are given by

(4a) $$\partial v/\partial C_G = \partial g/\partial C_G\, p(C_W, S_i)$$

and

(4b) $$\partial v/\partial C_W = \partial g/\partial C_W\, p(C_W, S_i) + \partial p/\partial C_W\, g(C_G, C_W, S_i)$$

The partial derivative given in (4a) is unambiguously negative. That is, the unconditional probability of visitation slopes downward with the cost of visiting the site. The sign of (4b) is ambiguous. The first term represents substitution between the Gauley River and other whitewater sites, and is positive. The second term represents the influence of access to whitewater on the probability of learning to paddle, and is negative. We refer to this second term as the "learning" effect, in that you must learn to paddle before you can visit the Gauley River as a private paddler. Thus, whether increased availability of whitewater rivers increases or decreases the unconditional probability of visitation depends on whether the substitution effect or the learning effect dominates.

In contrast, for a commercial rafter, previous experience and learned skills are not required to paddle the Gauley River. There is therefore no learning effect, which is equivalent to having $\partial p/\partial C_W = 0$. For commercial rafters, therefore, the sign of (4b) is unambiguously positive. Thus, for commercial paddlers, $\partial v/\partial C_W$ should be positive, due to the substitution effect, but for private paddlers, the sign of $\partial v/\partial C_W$ is indeterminate.

Ideally, data on individual decisions whether to learn to paddle, and then whether to visit the Gauley River on a specific recreation occasion, would allow direct estimation of p( ), g( ), and v( ), using a two-stage individual observations random utility model. Often, however, data limitations require estimation of an aggregated, zonal model. While such data will not allow direct identification of p( ) or of g( ), it can be used to estimate v( ). The results regarding $\partial v/\partial C_W$ apply to a zonal model as well, as differences in the individual probabilities aggregate to differences in rates of visitation at the zonal level. The influence of individual socio-demographic characteristics also aggregate to a zonal model, using aggregate measures of socio-demographics, though some information is lost in the aggregation (Hellerstein, 1995).

Regardless of whether an individual observations model or an aggregated model is used, it is important to include C_W as a demand shifter when estimating the reduced form model v( ). If C_W and C_G are spatially correlated, failure to include C_W will bias the estimated parameter on C_G, and therefore bias the estimated consumer surplus generated by the site.

Modeling Demand for Trips to the Gauley River

Summersville Lake is drawn down each fall to create capacity to absorb winter and spring rains. The water releases associated with these draw downs are usually done during daylight hours to encourage whitewater recreation, with flows set at appropriate levels for recreation. Recreational releases typically occur over six weekends in September and October, with additional releases on many intervening weekdays. Most visitation occurs on the weekends. The typical visit involves driving to the site on Thursday or Friday night, paddling for one, two or three days, and then driving home. Releases also occur on Mondays, but use rates are much lower (National Park Service).

Data

On three consecutive Saturdays in September 1991, the National Park Service conducted a census of all paddlers on this stretch of the Gauley River. The purpose of the census was to characterize use patterns and assess the need for developed access points along the river. Useful for our purpose was a question asking for the home zip code of each paddler. No demographic information was obtained in the survey, and no attempt was made to account for multiple visits made over more than one weekend, so estimation of an individual observation model was not feasible. Instead, zonal travel cost models were estimated for private paddlers and for commercial rafters.

The origin information was used to construct measures of visitation from all 2049 counties located in states east of or bordering the Mississippi River[1]. 8992 commercial rafters and 1531 private paddlers gave information on home zip code. These were aggregated to county level. An additional 56 paddlers came from origins outside the defined market area, including several from foreign countries. These were excluded from the analysis. Due to sampling difficulties, zip codes were not obtained from all paddlers. The National Park Service counted a total of 14002 paddlers on these three dates. Of these, 1878 were private paddlers (National Park Service). The commercial count included rafting company employees (guides, trainees, and video camera operators), but these were not surveyed and did not give zip code information. A separate study by Marshall University found that company employees accounted for 13.7% of commercial paddlers on these dates, which would imply that 1673 of the NPS-identified commercial paddlers were employees. The average number of private paddlers and paying commercial rafters on each of these three Saturdays was then 626.0 and 3483.7, respectively.

Information is not available on length of stay at the river. It was assumed that every weekend visitor paddled on Saturday. To the extent that visitors paddle only on Friday or only on Sunday, this assumption will underestimate total visitation. Both private and commercial paddler counts are uniformly higher on Saturdays than on either Fridays or Sundays (National Park Service), indicating that of those who paddle only one day, more paddle only on Saturday than only on Friday or only on Sunday. Average visitation per weekend was calculated for each county by multiplying the number of identified visitors from that county by 626.0/1531 for private paddlers and by 3483.7/8992 for commercial paddlers. These counts were converted to visitation rates by dividing by the total population of the county between the ages of 15 and 59, as reported in the 1990 Census. This age range includes the vast majority of whitewater paddlers.

Travel distance and time to the Gauley River were calculated for each county of origin using a route-finding software package. For each county, the point of origin was defined as the main post office in the largest city in the county. The destination point was defined as the main post office in Swiss, WV, the downstream terminus of the river segment. The algorithm used by the software package chose a route that minimized travel distance, with some preference given to staying on major highways.

Vehicle operating costs were set at $0.115 per mile, the 1991 estimate of costs for fuel, oil, tires, and maintenance for an intermediate-sized car (US Department of Transportation). A second component of variable operation cost that is often neglected in recreation demand studies is vehicle depreciation. Here vehicle depreciation was estimated based on the marginal change in car value associated with extra miles driven. In September 1991, a four year old (1988 model) intermediate-sized car with 10,000 excess miles was worth $800 less than a similar car with average mileage (NADA), implying a marginal depreciation rate of $0.08 per mile. Total cost per mile driven was therefore $0.195. This cost was divided among passengers in the car. Based on non-random observations at the site and on intercept surveys at other U.S. Army Corps of Engineers recreation sites, it was assumed that, on average, each car held 2.5 paddlers.

The functional form chosen, which is discussed below, is linear in the explanatory variables. Fixed trip costs that do not vary across zones of origin therefore do not influence parameter estimates, other than the intercept, and do not affect estimates of consumer surplus. We therefore ignore fixed trip costs such as food and lodging and fees paid to commercial raft companies. If some of these costs vary systematically across zones of origin, then our estimated slope parameter can be biased. Unfortunately, the on-site survey did not collect information on trip expenditures.

The opportunity cost of travel time was assumed to be some fixed proportion, ω , of the average wage rate in the county of origin (following Cesario and Knetsch). Average wage rates were calculated by dividing total wage income for the county by the number of employable persons aged 16 or older, and then dividing by 2000 hours per year. Wage income and number of employable persons were from the 1990 U.S. Census. Clear guidance does not exist for choosing a value of ω. Recent studies have used values of ω in the range of 0.25 (Needelman and Kealy) to 0.333 (Loomis et al.). Instead of imposing a value, we use the value that provides the best statistical fit. This is possible because travel time was estimated using routing-finding software, instead of as a constant multiple of travel distance. While travel time was highly correlated with travel distance, the correlation was not perfect.

In the conceptual model presented earlier, C_W represented the cost of accessing whitewater rivers. This cost is difficult to measure in practice, as it will include costs of visiting an array of sites. In principle, C_W will be lower for individuals who live in an area close to many whitewater rivers, and higher for individuals who live in an area that has fewer rivers located further away. We operationalize this concept by constructing an index of the availability of whitewater recreation, similar to indices used by Mullen and Menz for fishing sites, and by Loomis et al. for reservoirs. The Nationwide Whitewater Inventory (Barrow) lists all whitewater rivers in the U.S., including information on length, class of difficulty, and location (county). We included all river segments in the eastern U.S. of Class II (Novice) through Class V (expert). For county k, the specific form of the whitewater availability index was

(5) $$WAI_k = \sum_{j=1}^{J} \frac{L_j}{D_{kj}}$$

where L_j is the total length of all rivers located in county j and D_{kj} is the distance in miles between county k and county j as estimated by route-finding software. Summation is over all J=2049 counties in the eastern U.S. The index has the properties that a river is more important if it is closer, and more important if it is longer in length. Thus, WAI is inversely related to C_W. We would therefore expect the demand model for commercial rafters to have a negative coefficient on WAI, while the coefficient for private paddlers is indeterminate.

Values of WAI ranged from 6.2 to 44.4. The geographic distribution of high and low values is shown in Figure 1. As would be expected, higher values tended to occur in the more mountainous areas, which included the Appalachians and New England. Isolated counties with high values in the upper Midwest and southwest Missouri are associated with rivers draining into Lake Superior and rivers in the Ozarks, respectively. The rest of the Midwest and the deep south, with flatter topography, had fewer nearby whitewater opportunities. An important consequence of this spatial distribution is that values of WAI are strongly correlated (negatively) with distance from the Gauley River (ρ = -0.70). This correlation will have important implications in the model estimation.

We do not include any measure of availability of substitute activities other than whitewater recreation. For both private paddlers and commercial rafters, there are of course many other substitute activities that compete for recreation time (hiking, cycling, etc.). However, there is no particular reason to suspect that the spatial distribution of sites suitable for those activities is correlated with the location of whitewater rivers, so failure to include non-paddling substitute opportunities in a statistical model should not introduce omitted variable bias.

Finally, per capita income and population density in the county of origin, as reported in the 1990 Census, were included as demand shifters. If trips to the Gauley River are a normal good, participation rates should be higher from counties with higher per capita incomes. We have no prior expectation for influence of population density, which serves as a proxy measure of the degree of urbanization of the home county.

Statistical Model

With information on only 10523 visitors from 2049 counties, many counties had visitation rates of zero (68% of counties sent no commercial rafters, 84% sent no private paddlers). To accommodate these zero observations, we used a modified Heckman model, with a log-transformed dependent variable in the second stage regression. The Heckman model was estimated in two stages, with the first stage modeling the probability that positive visitation will be observed from a given county, and the second stage modeling the expected rate of visitation, given that positive visits are observed.

The average number of visitors per weekend from county i is denoted as N_i. Population in county i between the ages of 15 and 59 is denoted as P_i. The probability that county i had a positive number of visitors was modeled using a probit regression on independent variables X_i, so that

(6) $$N_i > 0 \text{ if } t_i > 0,$$

$$\text{where } t_i = X_i\beta + \mu_i \text{ and } \mu_i \sim N(0,\sigma^2_\mu).$$

Because this model is overparameterized, the variance, σ^2_μ is normalized to 1. The probability that positive visits will be observed from county i is then

(7) $$\Pr(N_i > 0) = \Phi(X_i\beta)$$

where $\Phi(\,)$ is the standard normal cumulative density function. For each county with positive visitation, the estimated parameter vector was used to calculate the inverse Mills ratio, defined as $\phi(X_i\beta)/\Phi(X_i\beta)$, where $\phi(\,)$ is the standard normal probability density function.

In the second stage, log-transformed visitation rates were regressed on independent variables Y_i, so that

(8) $$\ln(N_i/P_i) = Y_i\,\gamma + \nu_i$$

with $\nu_i \sim N(0, \sigma^2_v)$. The Inverse Mills Ratio was included in Y_i to account for possible correlation between μ_i and ν_i. This form is mostly similar to that discussed by Bockstael et al. (1990), but differs in that the left hand side of the second stage regression is log-transformed. This transformation provided better statistical fits to the data than an untransformed model, and avoids the theoretical problem that can occur when the random variate ν_i is large and negative. In such cases, an untransformed conditional visitation function would predict negative visits. Transformation assures that the conditional visitation function is always positive. With this transformation, the conditional visitation rate has a log-normal distribution, with mean

$$E(N_i | N_i > 0) = P_i \exp(Y_i \gamma + \sigma^2_v / 2) \tag{9}$$

(Hastings and Peacock 1975).

We stress that the empirical model is estimated in two stages to accommodate the large number of zones with zero visits. The two-stage model does not represent the two decisions faced by an individual. Both stages of the empirical model, taken together, estimate the aggregated unconditional probability of visitation, v( ).

For an individual county, the expected consumer surplus associated with visitation to the Gauley River is given by

$$E(CS) = \int_{TC_i}^{\infty} \Pr(N_i > 0) E(N_i | N_i > 0) dTC = \int_{TC_i}^{\infty} \Phi(X_i\beta) P_i \exp(Y_i\gamma + \frac{\sigma^2_v}{2}) dTC \tag{10}$$

where TC_i is the travel cost, including opportunity cost of time and fixed costs, from county i to the Gauley River. In our models, travel costs were included as an explanatory variable in both stages of the model. TC therefore appears in (10) in both X_i and Y_i directly, and again in Y_i through the inverse Mills ratio. This integral was evaluated numerically, with an upper limit of integration of $276, the largest calculated travel cost in the data set.

Bockstael et al. argue that if visitation rates are measured accurately, and the objective is to calculate consumer surplus for past trips, then the observed number of trips taken can be used in (10) instead of the predicted number. For this functional form, that approach yields a very simple formula for average consumer surplus per trip, namely average CS per trip = $-1/\gamma_{TC}$, where γ_{TC} is the coefficient on travel cost in the second stage regression. It is not clear, however, whether that approach is valid when travel costs are included in X_i. Regardless, for our data, the two approaches gave very similar estimates of average CS per trip. Results for both methods of calculation are presented.

Results

Through a search, values of ω were found that minimized the log-likelihood of both stages of the model, for both private and commercial paddlers. This approach is similar in spirit to that of McConnell and Strand. A likelihood ratio test showed that the estimated value of ω did not differ significantly between user groups ($\alpha > 0.30$). A single value of ω was therefore found that minimized the combined log-likelihood of both user group models. This single estimate of ω=0.20 is significantly different from 0 ($\alpha < 0.01$) and is significantly different from 0.5 ($\alpha < 0.01$), but is not significantly different from 0.25 ($\alpha > 0.50$) or from 0.333 ($\alpha > 0.10$)

Table 1 shows the first and second stage coefficient estimates for both private and commercial paddlers (columns 1 and 2), under the assumption that ω=0.20. All estimated coefficients on independent variables are significantly different from zero at the 5% confidence level, and most are significant at the 1% level. Population is included as an independent variable in the first stage, as more populous counties are more likely to send visitors. In the second stage, population is incorporated into the dependent variable. The own-price coefficients on travel cost are all negative, as expected. The coefficients on density were all negative as well, suggesting that proportionally fewer urban residents participate in whitewater recreation than suburban and rural residents. Coefficients on income are all positive, suggesting that whitewater paddling is a normal good.

As expected, higher values of the whitewater availability index led to lower visitation rates for commercial rafters. This result held for both stages of the model. For private paddlers, the relationship was reversed - higher values of the whitewater availability index led to higher visitation rates. This result also held for both stages of the model. Apparently, the learning effect dominates the substitution effect for private paddlers. Table 2 also presents estimates of average consumer surplus per trip, both using numerical integration of predicted number of trips, and the analytic solution based on observed number of trips.

To demonstrate the importance of accounting for access to other rivers when modeling visitation to a site such as the Gauley River, and the differential impact of the potential omitted variable bias across the two user groups, the same models were estimated without the whitewater availability index as a demand shifter. Estimation results are shown in columns 3 and 4 of Table 1. Because WAI is correlated with travel cost, its omission biases the parameter estimates on travel cost, and the estimates of consumer surplus. For private paddlers, this omitted variable bias makes the slope parameter estimates more negative, and therefore pushes the consumer surplus estimate down. For commercial rafters, the omitted variable bias makes the slope parameter estimates less negative, and pushes the consumer surplus estimate up.

Conclusions and Discussion

As has been shown in previous studies, failure to consider substitute sites can bias estimates of demand elasticities, and therefore bias estimates of consumer surplus. For commercial rafters, failure to model availability of substitute rivers (through inclusion of WAI) led to an average consumer surplus estimate that was 34% higher than the estimate obtained with the appropriately specified model. This result is consistent with that of Caulkins, Bishop and Bouwes and of Rosenthal.

What is new with this study is the demonstration that access to other recreation sites can influence visitation in the opposite direction, through the learning effect. Estimation of the private paddler model without the whitewater availability index results in an average consumer surplus estimate that was 34% *lower* than the estimate obtained with the appropriately specified model, and quite close to the estimate for commercial rafters. Omission of the whitewater availability index biases the slope parameter in opposite directions for these two groups, masking differences in consumer surplus.

The per trip estimates of consumer surplus obtained here are comparable to previous estimates. Bergstrom and Cordell estimated consumer surplus per trip at $30.66 for rafting and tubing and $20.66 for canoeing and kayaking. English and Bowker obtained estimates of per trip surplus for commercial rafting in Northern Georgia that ranged from $5.71 to $127, depending on assumptions regarding functional form and mileage costs. Using the specification most similar to that used in this study, their estimated consumer surplus per trip was $16.92. Daubert and Young, in a contingent valuation study, estimated per day values in Colorado that varied with flow rate in the river, with a maximum value of $33.26. Johnson, Bregenzer and Shelby, also in a contingent valuation study, obtained estimates of mean willingness to pay for a permit for access to a controlled whitewater river in Oregon of $32.66 and 52.93, depending on the question format used. More recently, though, Bowker, English and Donovan provided higher estimates of per trip consumer surplus for rafting. Their results for a value of ω close to that used in this study ranged from $125 to $193. They speculate, however, that their estimates may be biased upward somewhat by the truncated nature of their sample.

On the technical issue of valuing travel time, our results provide evidence that recreationists do value travel time, but at a rate less than their full wage rate. We estimate that recreationists value travel time at a rate equal to 0.18 times their average wage rate, fairly close to the commonly used values of 0.25 and 0.333. Likelihood ratio tests showed that the true value of ω falls somewhere between 0 and 0.5, and we cannot reject either of the commonly used values of 0.25 and 0.333. Smith, Desvousges, and McGivney used McConnell and Strand's approach to estimate ω for 23 different sites, and obtained estimates that varied widely, with most either negative or greater than one. Our estimate may have more precision because of the diversity of road types leading to our site from different directions, allowing more independence between travel distance and travel time.

Several caveats should accompany the results presented here. The on-site user survey collected very limited information, forcing assumptions about travel costs. Of particular importance were assumptions about number of individuals traveling per vehicle, the cost per mile driven of operating a car, and that costs other than gasoline, vehicle wear, and the opportunity cost of time were constant across all visitors. As travel distance increases, there may be a tendency to increase the number of travelers per vehicle. At the same time, longer distances may increase other costs such as costs for food and lodging. Mispecification of these costs will bias the estimated parameter on the travel cost variable, biasing estimated consumer surplus estimates. However, these mispecifications should not cause the parameter on WAI to switch sign, so we are confident that those results are robust.

Second, the fall water releases at the Gauley River are somewhat unique in that they occur during a season when there is low rainfall, and natural flow levels in most rivers in the east are low. Thus, there are fewer whitewater substitutes available than during the wetter spring. Had sampling been done when more rivers were flowing at adequate levels, the substitute effect might have been stronger than the learning effect, and the observed coefficient on WAI might have been negative for private paddlers. Indeed, the Gauley River does receive much lower visitation during spring releases than during the fall, suggesting substitution to other rivers when flows are high everywhere. The consumer surplus estimates presented here will therefore be strictly valid only for fall releases.

Finally, we mention the confounding issue of "moving to the site." Particularly for activities that involve a commitment of time and resources to take up, participants my choose residence location based in part on resource availability. Thus, persons who have chosen to become whitewater paddlers will tend to locate in areas closer to whitewater rivers. Thus, we cannot conclude with certainty whether availability of whitewater sites induces nearby residents to take up the sport, or attracts existing participants to the area. Either process involves a prior decision that results in more paddlers living in areas with more whitewater rivers.

Participation rates in outdoor activities that require specialized skills and experience are increasing. Policy decisions regarding access for such activities at high profile sites such as Mount McKinley, Yosemite National Park, and the Grand Canyon require reliable estimates of consumer surplus values accruing to participants in these activities. To fully understand participation decisions at sites that require prior training and experience, it is necessary to also model prior decisions regarding participation in the sport. This study accomplishes that goal using the simple approach of including a cross price term in a zonal demand model, and demonstrates that prior decisions and learning can be very important. To fully understand the role of prior experience and learning would require a more detailed random utility model of individual behavior, with individual data on both the decision to learn the sport and the decision to visit the particular site of interest.

Footnotes

[1] The defined market area is larger than that typically used, raising the possibility that visitors from more distant counties were engaged in multi-purpose trips (Smith and Kopp). The survey did not inquire about other activities during the trip, but conversations at the site indicated that single-use visitors did come from as far away as Minnesota and Florida.

Table 1. Coefficient and consumer surplus estimates for Gauley River visitation models.

	Full Models		Models Without WAI	
	Private (1)	Commercial (2)	Private (3)	Commercial (4)
First Stage				
Constant	-3.359 (13.31)	-0.514 (1.93)	-1.957 (10.71)	-1.425 (7.42)
Travel Cost	-0.00810 (6.62)	-0.0244 (16.76)	-0.0140 (13.45)	-0.0195 (19.68)
WAI	0.0611 (8.37)	-0.0409 (4.92)		
Per Capita Income	0.150×10^{-3} (9.27)	0.240×10^{-3} (12.95)	0.174×10^{-3} (11.12)	0.219×10^{-3} (12.24)
Population Density	-0.086×10^{-3} (5.30)	-0.134×10^{-3} (3.81)	-0.080×10^{-3} (4.85)	-0.138×10^{-3} (4.72)
Population	0.352×10^{-5} (8.94)	0.891×10^{-5} (10.90)	0.345×10^{-5} (8.74)	0.867×10^{-5} (10.77)
Second Stage				
Constant	-7.369 (8.77)	-2.524 (8.23)	-5.866 (8.13)	-4.212 (13.85)
Travel Cost	-0.0179 (8.34)	-0.0374 (17.84)	-0.0291 (9.80)	-0.0279 (16.53)
WAI	0.0899 (6.58)	-0.0838 (9.10)		
Per Capita Income	0.099×10^{-3} (2.64)	0.242×10^{-3} (9.67)	0.173×10^{-3} (3.66)	0.190×10^{-3} (7.75)
Population Density	-0.574×10^{-4} (2.42)	-0.325×10^{-4} (2.32)	-0.595×10^{-4} (2.34)	-0.402×10^{-4} (2.81)
Inverse Mills Ratio	1.365 (5.86)	1.108 (8.67)	1.644 (6.10)	1.090 (7.98)
Consumer Surplus per trip				
Using predicted # trips	\$52.17	\$26.04	\$34.41	\$34.36
Using observed # trips	\$55.84	\$26.73	\$34.33	\$35.82

(t-values in parentheses)

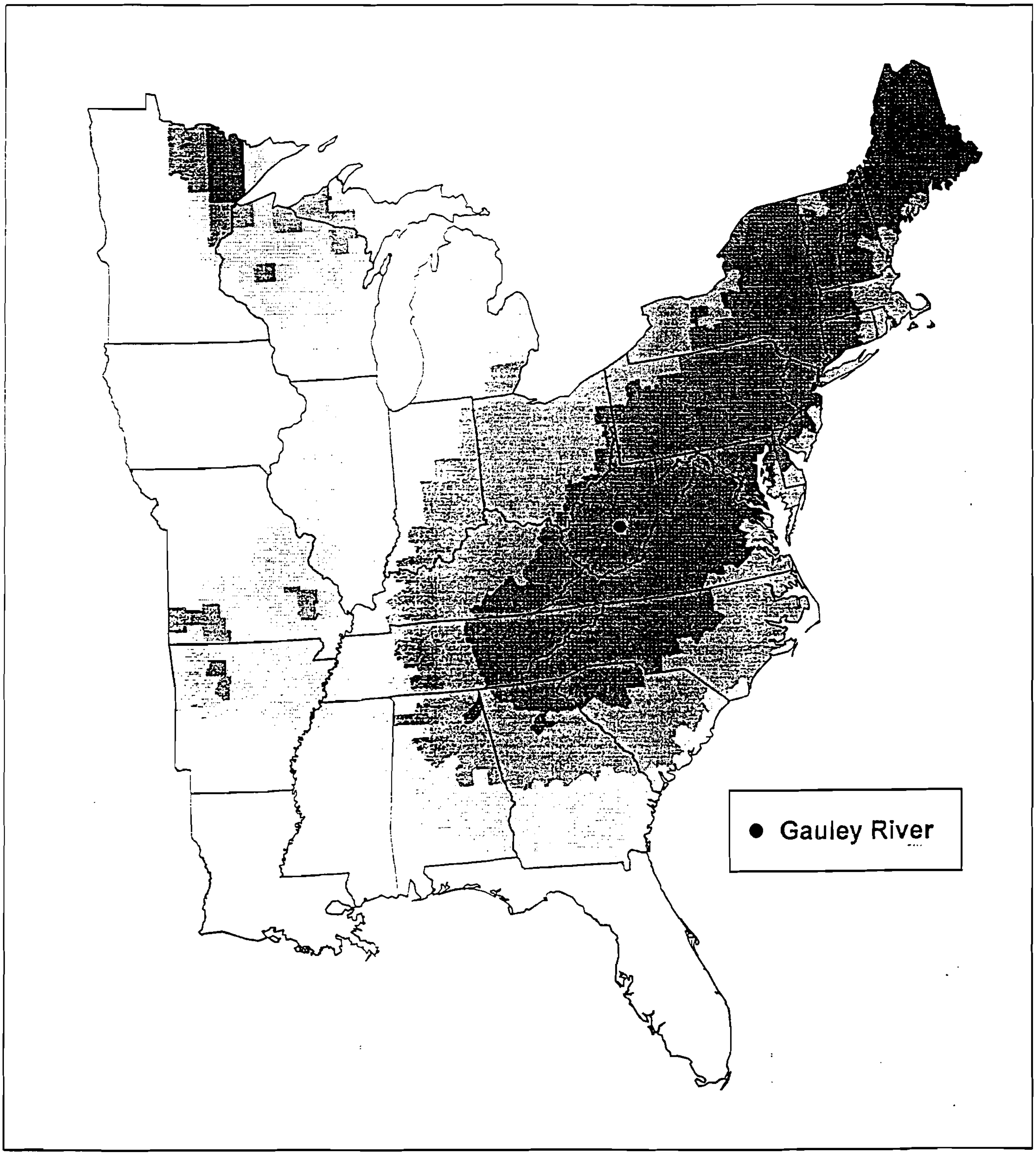


Figure 1. Quartiles for Whitewater Availability Index.

References

Barrow, P. 1990. **Nationwide Whitewater Inventory**, Second Edition. American Whitewater Affiliation, Washington, D.C.

Bergstrom, J.C., and H.K. Cordell. 1991. "An Analysis of the Demand for and Value of Outdoor Recreation in the United States." **Journal of Leisure Research** 23(1):67-86.

Bockstael, N.E., I.E. Strand Jr., K.E. McConnell, and Firuzeh Arsanjani. 1990. "Sample Selection Bias in the Estimation of Recreation Demand Functions: An Application to Sportfishing." **Land Economics** 66(1):40-49.

Bowker, J.M., D. K. English, and J.A. Donovan. 1996. "Toward a Value for Guided Rafting on Southern Rivers." **Journal of Agricultural and Applied Economics** 28(2):423-432.

Burrell, B., and P. Davidson. 1985. **A Wildwater West Virginia**

Caulkins, P.B., R.C. Bishop, and N.W. Bouwes. 1985. "Omitted Cross Price Variable Biases in the Linear Travel Cost Model: Correcting Common Misconceptions." **Land Economics** 61(2):182-187.

Cesario, F.J., and J.L. Knetsch. 1976. "A Recreation Site Demand and Benefit Estimation Model." **Regional Studies** 10:97-104.

Daubert, J.T., and R.A. Young. 1981. "Recreational Demands for Maintaining In-stream Flows: A Contingent Valuation Approach." **American Journal of Agricultural Economics** 63(4):666-676.

English, D.B.K., and J.M. Bowker. 1995. "Sensitivity of Whitewater Rafting Consumer Surplus To Pecuniary Travel Cost Specifications." Paper presented at annual meetings of the Southern Agricultural Economics Association, New Orleans, LA, February 1995.

Johnson, R.L., N.S. Bregenzer, and B. Shelby. 1990. "Contingent Valuation Question Formats: Dichotomous Choice versus Open-Ended Responses" in R.L. Johnson and G.V. Johnson eds. **Economic Valuation of Natural Resources: Issues, Theory, and Applications**. Westview Press, Boulder Co.

Hastings, N.A.J., and J.B. Peacock. 1975. **Statistical Distributions: A Handbook for Students and Practitioners**. John Wiley & Sons, New York.

Hellerstein, D. 1995. "Welfare Estimation Using Aggregate and Individual-Observation Models: A Comparison Using Monte Carlo Techniques." **American Journal of Agricultural Economics**. 77(3):620-630.

Loomis, J., B. Roach, F. Ward, and R. Ready. 1995. "Testing Transferability of Recreation Demand Models Across Regions: A Study of Corps of Engineers Reservoirs." **Water Resources Research** 31(3):721-730.

McConnell, K.E., and I. Strand. 1981. "Measuring the Cost of Time in Recreational Demand Analysis: An Application to Sport Fishing." **American Journal of Agricultural Economics** 63(1):153-156.

Mullen, J.K., and F.C. Menz. 1985. "The Effects of Acidification Damages on the Economic Value of the Adirondack Fishery to New York Anglers." **American Journal of Agricultural Economics** 67(1):112-119.

NADA. 1991. **Official Used Car Guide** Vol 58 No 9 (September).

National Park Service. 1992. "Survey of Whitewater Boating Use on the Gauley River: Final Report" Gauley River National Recreation Area, Glen Jean, WV.

Needelman, M.S., and M.J. Kealy. 1995. "Recreational Swimming Benefits of New Hampshire Lake Water Quality Policies: An Application of a Repeated Discrete Choice Model." **Agricultural and Resource Economics Review** 24(1):78-87.

Rosenthal, D.H. 1987. "The Necessity for Substitute Prices in Recreation Demand Analysis." **American Journal of Agricultural Economics** 69(4):828-837.

Smith, V.K., W.H. Desvousges and M.P. McGivney. 1983. "The Opportunity Cost of Travel Time in Recreation Demand Models." **Land Economics** 59( ):259-277.

Smith, V. K., and R. J. Kopp. 1980. "The Spatial Limits of the Travel Cost Model." **Land Economics** 56(1):64-71.

U.S. Department of Transportation. 1994. **National Transportation Statistics - Annual Report, September 1993** U.S. Government Printing Office, Washington D.C.