Empirical Discrete/Continuous Choice Modeling for the Valuation of Non-market Resources or Public Goods

by

Trudy Ann Cameron
Department of Economics
University of California, Los Angeles

UCLA Working Paper No. 503
September 1988

Empirical Discrete/Continuous Choice Modeling
for the Valuation of Non-market Resources or Public Goods

by

Trudy Ann Cameron
Department of Economics
University of California, Los Angeles

ABSTRACT

Contingent valuation (CV) survey methods are now being used quite widely to assess the economic value of non-market resources and public goods. However, being "hypothetical responses to hypothetical questions," the implications of these surveys have sometimes met with a degree of skepticism. Here, hypothetical CV data are combined with "travel cost" data on actual market behavior (exhibited by the same consumers) to "auto-validate" the implied CV demand functions. Conditional upon the functional form chosen to represent the direct utility function of respondents (quadratic, in this case), it is possible to estimate *jointly* both the parameters of the underlying utility function and its corresponding Marshallian demand function. Equivalence of the utility parameters implied by the two types of data can be tested statistically; if the CV data are considered *a priori* to be less reliable, their influence in the joint estimation process can be down-weighted. Partial auto-validation can be achieved even if actual access days and income are the only additional data available. Respondent heterogeneity can be readily accommodated by allowing the utility parameters to vary with the levels of sociodemographic variables. A sample of Texas recreational anglers illustrates the technique. The calibrated utility functions are used to compute actual individual measures of access values in the form of empirical equivalent and compensating variations for various access restrictions.

* This project has benefited from helpful comments and suggestions provided by E.E. Leamer, W.M. Hanemann, J. Hirshleifer and K.E. McConnell. Jerry Clark and the Texas Department of Parks and Wildlife generously provided the responses to the 1987 creel and contingent valuation survey of the Texas Gulf Coast (designed in consultation with John Stoll of Texas A&M). Maury Osborn at Texas Parks and Wildlife cleaned up the data and provided a final data tape. The Inter-university Consortium for Social and Political Research provided Census data by 5-digit zip code for Texas. This research is supported in part by EPA cooperative agreement # CR-814656-01-0.

September 7, 1988

Empirical Discrete/Continuous Choice Modeling for the Valuation of Non-market Resources or Public Goods

1. INTRODUCTION

Despite the potential for a variety of biases in poorly designed contingent valuation (CV) surveys (described in detail by Cummings, Brookshire, and Schulze, 1986, and by Mitchell and Carson, 1988) there are still many situations where market simulations or actual market experiments are prohibitively difficult. CV surveys, designed to elicit respondents' reactions to pricing scenarios in hypothetical market situations, can provide a valuable source of information about the characteristics of demand for a good not presently priced and traded in a real market. The technique has now been widely applied to the valuation of non-market resources and public goods. Interpreted cautiously, the findings generated by models using appropriate CV data can sometimes be used to produce "ballpark" estimates of the collective social value of access to the non-market good.

CV data has been criticized as "hypothetical answers to hypothetical questions." Consequently, "external validation" of empirical applications of CV methods has received considerable attention in the applied literature. Bishop and Heberlein (1979) and Bishop, Heberlein and Kealy (1983) question the ability of contingent valuation survey methods to elicit estimates of resource values consistent with other estimates derived via the travel cost methods or from actual cash transactions in simulated markets. They conclude that CV mechanisms produce "meaningful--albeit inaccurate--economic information."

Similar conclusions are reached by Schulze, d'Arge, and Brookshire (1981) in their review of research using different methods for valuing non-market environmental commodities. They determine that "all evidence obtained to date suggests that the most readily applicable methodologies for evaluating environmental quality--hedonic studies of property values or wages, travel cost, and [CV] survey techniques--all yield values well within one order of magnitude in accuracy. Such information...is preferable to complete ignorance."

Brookshire, Thayer, Schulze, and D'Arge (1982) also seek external validation of hypothetical methods. They compare CV estimates of the value of a public good with corresponding estimates derived from a hedonic property value study, finding that survey responses yield estimates of willingness to pay for a public good which are consistent with a hedonic-market analysis. Regarding CV survey techniques for valuation, they conclude that "[a]lthough better accuracy would be highly desirable, in many cases where no other technique is available for valuing public goods, this level of accuracy is certainly preferable to no information for the decision-making process." Sellar, Stoll and Chavas (1985) compare value estimates from two different CV methods and the travel cost approach. They conclude that the "referendum" (or "closed-ended") form of the contingent valuation method and the travel cost method do provide comparable estimates of consumer surplus, and that whenever possible, *both* methods should be used in future studies as a validity check on the results.

Brookshire and Coursey (1987) compare hypothetical non-market CV responses with market-like elicitation processes (Vernon Smith's public good auction experiments in the laboratory and in the field). They examine the so-called "loss-aversion phenomenon" (i.e. the case where the compensating variation vastly exceeds the equivalent variation for the same change) which seems particularly pronounced in CV methods. They find that even a small number of repetitions of the auction process dramatically reduce (although do not eliminate) the evidence of loss-aversion behavior. Compared to CV methods, the marketplace appears to be "a strong disciplinarian" in terms of limiting loss-aversion tendencies.

In all these previous studies aimed at external validation of the values for non-market goods produced by the CV method, the alternative measures of value were obtained either by *indirect* methods (the travel cost approach or hedonic wage or rent functions) or by small *simulated market* experiments. The models introduced in this paper will also appeal to the marketplace to discipline contingent valuation estimates, but the validation will occur in the context of a *single* model applied to a *single* sample. Hence the term "auto-validation." The reinforcement goes both ways: the observed actual market behavior of individuals "disciplines" the professed implicit demand information extracted from contingent valuation responses. At the same time, the CV information provides insights into the probable behavior of respondents under conditions which are far removed from the current market scenario. These new models incorporate *both* the hypothetical *discrete choice* elicited by the CV question and the theoretically consistent *continuous* model of consumer demand. The framework is similar to that employed by Hanemann (1984) and by Dubin and McFadden (1984), but in the present case, the discrete choices are purely hypothetical.

CV methods are applicable for valuing a wide variety of non-market, public, and pre-test-market goods. (Mitchell and Carson, 1988, document well over 100 studies.) For concreteness, this research concentrates on the demand for access to a recreational fishery. Recreational fisheries valuation has attracted considerable policy-making interest over the past few years, and a wide variety of theoretical examinations and empirical attempts at valuation are extant. (To cite only a few of the more recent studies: McConnell, 1979, Anderson, 1980, 1983, Samples and Bishop, 1980, McConnell and Strand, 1981, Vaughn and Russell, 1982, Morey and Rowe, 1985, Rowe, Morey, Ross, and Shaw, 1985, Samples and Bishop, 1985, Donnelly, Loomis, Sorg, and Nelson, 1985, Morey and Shaw, 1986, Cameron and James, 1986, 1987, Huppert and Thomson, 1987, Cameron 1988, Cameron and Huppert, 1988a,b,c, Agnello, 1988, and McConnell and Norton, undated). Recreational fisheries valuation research will also assume increasing importance as quantification of the social costs of acid precipitation proceeds. The U.S. Fish and Wildlife Service estimates that economic activity associated with recreational fishing generated $17.3 billion in 1980 and $28.1 billion in 1985. There are at least 60 million Americans who fish regularly.[1] Acid rain kills fish, which undeniably affects the consumer surplus associated with recreational fishing.

Section 2 of this paper develops the logic whereby a direct *utility difference* function (defined over fishing days and a composite of all other goods) and the corresponding Marshallian demand function for fishing access days can be modeled jointly. Section 3 describes a sample of CV and travel cost data from a larger survey used to demonstrate this technique. Section 4 describes alternative stochastic specifications. Section 5 provides a general outline of the types of results these models generate. Section 6 goes into detail regarding the specific empirical results for a basic model and several extensions. Considerations for ensuring transferability of these methods to other applications are addressed in Section 7. Section 8 concludes with several important caveats.

2. CONTINGENT VALUATION QUESTIONS AND UTILITY DIFFERENCE FUNCTIONS

Background

Recent research by McConnell (1988) has built upon the utility-theoretic tradition in the analysis of discrete-choice contingent valuation data initiated by Hanemann (1984), which grew from early path-breaking work by Bishop and Heberlein (1979). Utility-theoretic models should be the econometric objective whenever such models are feasible. However, data deficiencies and empirical expedience will often dictate that the ultimate specification of a valuation function for a non-market resource take a simple and arbitrary ("atheoretic") form. These forms can, at best, be interpreted as "approximations to some underlying utility-theoretic inverse demand function." Nevertheless, naive specifications can be quite successful empirically and are widely used in the analysis of data in all areas of economics.

One popular format for CV surveys is the "referendum" question. It is often argued to be less subject to some of the usual CV biases than are other formats. Rather than asking the respondent to place his own specific dollar value on access to the resource, the respondent is offered a single threshold value and asked to indicate whether his personal valuation is greater or less than this amount. For referendum CV survey data, simple reinterpretations of the estimated coefficients generated by conventional maximum likelihood probit or logit models for these discrete responses can produce fitted "valuation" functions quite readily (See Cameron and James, 1987, and Cameron, 1988). But the majority of empirical examples thus far have been pragmatic, rather than utility theoretic. Nevertheless, considerable information can be extracted about the basic relationship between (a.) whatever theoretical concept of value is elicited by the wording of the CV question and (b.) any explanatory variables which might have been collected by the survey.

Now that simple and tractable estimation procedures have been developed for arbitrary specifications which use referendum CV data, research attention should certainly be turned to the development of *estimating* specifications that are consistent with the economic theory of constrained utility maximization. When an implicit Marshallian demand function for a non-market good is derived from some specific underlying utility specification, one can draw upon microeconomic theory to generate testable hypotheses regarding the signs and sizes of estimated coefficients. Without this structure, any set of "demand" coefficients is as good as another.

For the survey available for this study, there are unfortunately *two* possible ways in which the referendum CV question posed to the respondent could be interpreted. Respondents are asked if they would quit using the resource entirely if the cost of access had been a certain number of dollars (T) higher.

In the first instance, the question could be construed as asking whether the respondent would entirely cease to use the resource if the annual access fee ("tax") were equal to T. Let Y be the respondent's income, let q^0 be the current number of trips per year to the recreation site, and let M be the respondent's typical market cost of access (travel) and incidental expenses on complementary market goods associated with one trip.

With cross-sectional data, it is convenient to assume that for similar individuals, the common utility function is such that access to the recreational resource is *separable* from a composite of all other goods and services, z, for which the price can be normalized to unity. The market price of access, M, is approximately the "travel cost" used in many earlier studies to extract properties of the the demand function for access.[2] If it can be assumed that these complementary goods are consumed *in fixed proportions* with the number of recreation trips, then only the number of trips appears separately in the utility function: U(z, q). Market goods are purchased only if access is chosen; if the individual decides not to consume any access days, his expenditures on these complementary goods will be zero as well.[3]

If the respondent indicates that he would continue fishing despite the tax T, this implies that his utility when paying the tax and enjoying access exceeds his utility when forgoing all trips and thereby avoiding the tax (and also the incidental expenditures and travel costs associated with each trip):

$$U(Y - Mq^1 - T, q^1) > U(Y, 0). \tag{1}$$

Crucially, as pointed out by McConnell (1988), the post-tax quantity demanded is *endogenously* determined in this model. Observed current consumption of access days, q^0, will probably be larger than the unobserved subsequent "post-tax" consumption, q^1, if income effects are present. In a deterministic world, one would expect to see individuals state that they would *not* cease to use the resource if their utility difference is positive:

$$\Delta U = U(Y - Mq^1 - T, q^1) - U(Y, 0) > 0. \tag{2}$$

However, without additional information on the demand function for access days that will allow q^1 to be identified, any utility-theoretic discrete choice model for referendum data will be difficult to formulate.

But an alternative interpretation of the CV question is possible. Perhaps respondents think of the access fee T as implicitly reflecting a *price* change at their current consumption level, q^0, rather than a lump sum tax.

They may interpret the question as asking whether or not they would choose non-zero access days if the price per day went from M to $M+(T/q^0)$. Rather than a *parallel shift* of the budget constraint, then, we might conceivable have the CV question interpreted as suggesting a *change in the slope* of the budget constraint. In this case, the the CV question would seem to be asking respondents whether their post-price change optimal consumption of access days would be positive. (I.e. if their optimal number of access days was negative, their highest utility would correspond to zero access days, providing that preferences are well-behaved.) Rather than the *utility-difference* underlying the discrete response in equation (2), this *projected optimal consumption level* would "drive" the discrete choice portion of the model. The results reported in this paper will emphasize the "lump sum tax" interpretation, but some results for the alternative "price change" interpretation will be provided for comparison, since the interpretation *does* affect the resulting estimates of resource value.

Since Hanemann (1984) first popularized utility-theoretic modeling of referendum CV responses, several researchers (e.g. Sellar, Stoll, and Chavas, 1983, 1987) have employed maximum likelihood discrete choice models (primarily logit models) with simple linear-in-parameters utility functions. Often, these functions are also linear-in-variables, which makes them *very* simple representations of utility. This research will employ a more general quadratic specification.

The conventional discrete choice approach uses the utility *difference* function, ΔU, which is linear in the *same* parameters as is U only if the underlying utility function is linear-in-parameters. The value of access to the resource (the equivalent variation for a loss of access) can be interpreted as the lump sum tax T (= T*) which would make someone just indifferent between paying that tax and continuing to use the resource, or forgoing both access and the tax. If the utility difference function ΔU is linear and additively separable in a univariate function of T, then it is algebraically simple to solve the fitted formula for ΔU to yield a point estimate for T^*. Under these special conditions, it is also possible to estimate *exactly* the same underlying model using the logistic or normal censored regression models for T^* proposed in Cameron (1988). As the utility function U becomes more complex, however, it can become much more difficult to solve the indifference equality to yield a simple expression for T^*. Furthermore, if one wishes to use packaged discrete choice models, specifications for U must be linear-in-parameters. Fortunately, due to empirical economists' history of reliance on ordinary least squares (OLS) based methods, several such families of linear-in-parameters direct utility or indirect utility functions are familiar.

But the CV question is purely hypothetical, and researchers and policy-makers are correct to exercise caution in utilizing the results produced by these methods. In this paper, the hypothetical market information extracted from respondents by the CV survey is combined with information regarding *actual* behavior--namely, current consumption of resource access when specific per-unit access prices (in excess of current travel costs) are zero. Requiring that observed market behavior be consistent with the *same* utility function being calibrated by the discrete choice estimation of the utility difference function, ΔU, can presumably ensure greater reliability of the results.

Candidate Specific Functional Forms for Utility

All of the variants of the basic model proposed in this paper are based on a *quadratic* direct utility specification. This is certainly one of the most algebraically simple choices. Its selection requires some justification.

Utility-theoretic discrete choice models in the broader literature usually work from specifications of *indirect* utility, since commodity prices (rather than quantities) are more plausibly assumed to be exogenous for the typical consumer. In this context, however, consumers are choosing between forgoing income T in order to gain access to the resource, or, alternatively, *keeping* the income and consuming *zero* access. Implicitly, choosing not to pay T implies that you face an infinite price for access, since access would then be prohibited.[4] Infinite prices are difficult to accommodate in *any* indirect utility function. Consequently, the choice model must be based on formulations of *direct* utility.

It is worth exploring why a number of familiar specifications other than the quadratic are unsuitable in this context. Other linear-in-parameters functions which have been widely used empirically include the translog and the generalized Leontief specifications. The translog is quadratic in the logarithms of the arguments:

$$(3) \qquad U(z,q) = \beta_1 \log z + \beta_2 \log q + 2\beta_3(\log z)^2 + \beta_4 (\log z)(\log q) + 2\beta_5 (\log q)^2 .$$

Here, however, it is critical that utility levels be *defined* and *non-zero* when consumption of one commodity (namely, recreation days) goes to zero. This disqualifies the ordinary translog model, since this function is only defined over strictly positive quantities of each good.[5]

The generalized Leontief specification satisfies the boundary requirements, and is generally considered to be a more "flexible" functional form than the quadratic:

$$(4) \qquad U(z,q) = 2\,\beta_1\, z^{1/2} + 2\,\beta_2\, q^{1/2} + \beta_3\, z + 2\,\beta_4\, z^{1/2}q^{1/2} + \beta_5\, q$$

However, while a generalized Leontief *indirect* utility function can readily be differentiated to yield Marshallian demands, this similar functional form for the *direct* utility function does not yield Marshallian demands as easily. Its associated Marshallian demand functions are *much* more complex than those corresponding to the quadratic utility function.

The Stone-Geary direct utility function has been widely used as a justification for the Linear Expenditure System (LES) of demand equations. It, too, is inappropriate for this particular application, again because it is necessary to have well-defined utility levels for zero consumption of access days. The empirical problems with the Stone-Geary model are outlined in Appendix I.

The quadratic form is a useful local approximation to any arbitrary surface. Why not then expand to third-order terms? Several of the quantities of interest which are derived from the calibrated model necessitate solving the fitted utility function for the value of one of its arguments. The standard formula for computing quadratic roots is straightforward to use. The formulas for the roots of cubic equations are considerably less easy. (See CRC, 1981, p.9.) However, subsequent empirical research should certainly explore such forms, especially if the results for quadratic utility specifications suggest that a higher degree of parameterization might be supported.

Contemporaneous work by Huppert (1988) employs an alternative strategy. He begins with a simple functional form (log-linear) for the Marshallian demand specification and accepts the corresponding (unnamed) functional form for the underlying utility function.[6]

The Quadratic Direct Utility Specification

For homogeneous consumers, the simplest quadratic utility in this context is:

$$U = \beta_1 z + \beta_2 q + \beta_3 z^2/2 + \beta_4 zq + \beta_5 q^2/2. \tag{5}$$

Under the current scenario for the respondent, consumption of the Hicksian composite good z is $(Y - Mq^0)$, just annual income minus total market expenditures associated with resource access, which will depend on the number of trips. Currently, q^0 will also be non-zero for everyone being interviewed.[7] For the hypothetical CV scenario, let U_T be expected utility *with* access to the resource *and* under an annual access tax of T. Let U_N be expected utility with no access (and hence no tax). Then, under the first interpretation of the CV question, the utility difference[8] which dictates a respondent's answer will be:

$$\begin{aligned} \Delta U = U_T - U_N = {} & \beta_1 [Y-Mq^1-T] + \beta_2 q^1 \\ & + \beta_3 [Y-Mq^1-T]^2/2 + \beta_4 [Y-Mq^1-T]q^1 + \beta_5 (q^1)^2/2 \\ & - \{ \beta_1 [Y] + \beta_3 [Y]^2/2 \}. \end{aligned} \tag{6}$$

Collecting terms involving the same parameters:

$$\begin{aligned} \Delta U = {} & \beta_1 \{[Y-Mq^1-T] - Y\} + \beta_2 q^1 \\ & + \beta_3 \{[Y-Mq^1-T]^2 - Y^2\}/2 + \beta_4 [Y-Mq^1-T]q^1 + \beta_5 (q^1)^2/2. \end{aligned} \tag{7}$$

If the utility surface adequately describes the configuration of individuals' preferences, then it should also be consistent with the current *observed* behavior, namely demand for access days in an environment where per-day access prices (beyond M) are currently zero. To this end, it is extremely useful to derive the Marshallian demand for *current* access days, q^0, associated with the utility function given in equation (3).

If some non-zero per-unit price, p, for access to recreation days were to be charged, the utility maximization hypothesis would suggest that the individual would implicitly choose their optimal z and q by maximizing the Lagrangian:

$$(8) \qquad L = \beta_1 z + \beta_2 q + \beta_3 z^2 + \beta_4 zq + \beta_5 q^2 + \lambda (Y - Mq - pq),$$

where λ is a Lagrange multiplier with the usual interpretation as the marginal utility of income. The first order conditions for this Lagrangian yield a corresponding Marshallian demand for q of:

$$(9) \qquad q = [\beta_2 + \beta_4 Y - \beta_1(M + p) - \beta_3 Y (M + p)] / [2(\beta_4)(M + p) - \beta_3 (M + p)^2 - \beta_5].$$

Since every additive term in both the numerator and denominator of this expression contains a multiplicative β coefficient, the demand function is invariant to the scale of the β vector. This is good, in that demands should be insensitive to the scale of measurement of utility. An infinite number of proportional β elements will satisfy the equality. Consequently, it is necessary to adopt some normalization of the demand function parameters (for example, $\beta_2 = 1$, an entirely arbitrary and inconsequential choice). Thus the form of the demand function actually estimated will be:

(10) $$q^0 = [\ 1 + (\beta_4/\beta_2)\ Y - (\beta_1/\beta_2)(M + p) - (\beta_3/\beta_2)\ Y\ (M + p)\ ]\ /$$
$$[\ 2(\beta_4/\beta_2)(M + p) - (\beta_3/\beta_2)\ (M + p)^2 - (\beta_5/\beta_2)\ ],$$

or, since p is identically zero in the sample used to estimate the model:

(11) $$q^0 = [\ 1 + (\beta_4{*})\ Y - (\beta_2{*})(M) - (\beta_3{*})\ Y\ (M)\ ]\ /$$
$$[\ 2(\beta_4{*})(M) - (\beta_3{*})\ (M)^2 - \beta_5{*}\ ].$$

The asterisks indicate that the estimable parameter is only identified relative to the value of β_2.

The derivatives with respect to p of the Marshallian demand function in equation (9) will be the same as the derivatives with respect to M, a property implicitly exploited by those who rely upon the "travel cost" method to value resources. However, this demand function is highly non-linear in p. A "naive" demand model for access days which is specified simply as an arbitrary linear-in-parameters function of M and Y will *not* yield parameter estimates regarding which neoclassical microeconomics can provide distinct null hypotheses. Nevertheless, non-linear least squares techniques could be used with equation (9). To the extent that the available data for M are typical (for that individual), and to the extent that these market goods and access days are always consumed in *fixed proportions*, the fitted parameters for this demand function will indeed reveal local approximations to (for example) the price and income elasticities of demand for access.

This research has two main objectives. First, it acknowledges the supposition that the utility maximizing decisions of economic agents, whether real *or* hypothetical, should reflect the *same* underlying structure of preferences. Conditional on the extent to which a quadratic function in z and q is an *adequate* representation of the preferences of individuals in this sample, this supposition will be used to impose parameter constraints across (a.) the *discrete choice model* used to explain yes/no responses to the *hypothetical* CV question and (b.) the *demand model* explaining actual *observed* levels of resource access under the zero-price scenario. Requiring that respondents' professed behavior in a hypothetical context be consistent with their observed behavior in real markets should attenuate the degree of bias due to the hypothetical nature of the CV question. In turn, the CV information allows the researcher to "fill in" some information about demand that is not captured within the range of the currently observable demand data.

A second objective is to make a rigorous statistical comparison of the different utility configurations implied by the CV and the travel cost data if the utility parameters implied by each are unconstrained. When parameter constraints are imposed across two models, it is often possible to allow the parameters to differ, taking on any values the data suggest. Contingent on the validity of the assumption of quadratic utility, it will then be interesting to test statistically the hypothesis that the corresponding parameters in the two models are the same. This is implicitly a test of whether professed behavior in the hypothetical market is consistent with observed behavior in a real market. If one "believes" the estimates of the demand function (based on actual markets), and if it is possible to *reject* the hypothesis of equality of the corresponding parameters in the discrete choice CV portion of the model, then it will be tempting to suspect that CV techniques might be unreliable. (Of course, such a hypothesis test is also implicitly a test of the adequacy of the quadratic specification, so the results will not necessarily be conclusive.[9]) However, in examples where it is *not possible* to reject parameter equality, this will be fairly strong support for use of the CV method!

In addition to these two substantial objectives, this paper describes in detail a number of extensions which demonstrate the flexibility of this model as a prototype for subsequent work.

3. AN ILLUSTRATIVE EXAMPLE: THE TEXAS PARKS AND WILDLIFE SURVEY

Between May and November of 1987, the Coastal Fisheries Branch of the Texas Department of Parks and Wildlife conducted a major creel survey of recreational fishermen from the Mexico border to the Louisiana state line. Vast quantities of information were collected on the details of the catch of each fishing party, but the "socioeconomic" portion of the survey is most pertinent here. The specific CV question asked of respondents was: "If the total cost of all your saltwater fishing last year was ____ more, would you have quit fishing completely?" At the start of each survey day, interviewers randomly chose a starting value from the list $50, $100, $200, $400, $600, $800, $1000, $1500, $5000, and $20,000. On each subsequent interview, the next value in the sequence was used. Therefore, offered values can be presumed to have no correlation whatsoever with the characteristics of any respondent. In addition to this question, respondents were asked "How much will you spend on this fishing trip from when you left home until you get home?" The survey also established how many trips the respondent made over the last year to all saltwater sites in Texas.[10] Five digit zip codes were collected, which allows establishment of residency in Texas.

Income data were not collected from each respondent, but the five-digit zip codes allows merging of the data with 1980 Census median household incomes for each zip code. While the use of group averages instead of individual income information undeniably involves errors-in-variables complications in the estimation process, the distortions may in fact be not much greater than they would be with the use of self-reported income data in an unofficial context. It is well known that many individuals have strong incentives to misrepresent their incomes if they do not perceive a legal requirement to state them correctly. Furthermore, zip codes cover relatively homogeneous "neighborhoods," at least when compared to income data on the county level, for example. Individuals' consumption patterns tend to conform somewhat to those of their neighbors, so median zip code income may be a better proxy for "permanent" disposable income than actual current self-reported income. There is high variance in median incomes across zip codes, so the Census income variable may actually make a substantial and accurate contribution to controlling for income heterogeneity among the survey respondents.

In other work utilizing the entire dataset (Cameron, Clark, and Stoll, 1988) it has been determined that subsets of individuals in the sample exhibit extreme behavior. The full sample has therefore been arbitrarily limited for this study. Since the initial models presume identical underlying utility functions for all individuals, who report more than sixty fishing trip per year are discarded from the sample. It is unlikely that these individuals are typical, since 90% of usable sample reports fewer than this number of days. The median number of trips reported is between eleven and twelve. This research is directed at "typical" anglers.

It is also the case in the full usable sample from the survey that some individuals respond that they would keep fishing even if the cost had been $20000 higher--when $20000 exceeds the median household income of their zip code! Since the assignment of value thresholds was completely exogenous, the estimating sample includes only those respondents who were posed values up to and including the $1500 offer.

The final criterion for inclusion in the sample for this study was that a respondent should not report spending more than $100 on this fishing trip.

Again, a very large proportion of the sample passes this criterion. When market expenditures are reported to be much larger than this, it seems reasonable to suspect that capital items have been included, so that it would be invalid to treat these costs as "typical" for a single fishing trip. Current expenditures over $2000 were reported by several respondents.

Descriptive statistics for the variables used in this paper are contained in Table I.

4. THE STOCHASTIC SPECIFICATION: JOINT ESTIMATION

In order to have the option of constraining the coefficients of the utility function and the corresponding Marshallian demand function to be identical, and in order to endogenize post-tax quantity q^1, the discrete choice model and the demand equation must be estimated simultaneously. The simplest assumption would be that the error terms associated with the decisions represented by these two models were statistically independent. However, it is highly probable that unmeasured attributes of an individual which make them more likely to choose to continue fishing under the hypothetical CV scenario also make their observed number of fishing days greater than the fitted demand function model suggests. To fix ideas, however, it is helpful to begin by considering the two parts of the model completely separately.

Unconstrained Parameter Estimates, Separate CV Choice Model

First of all, since only current optimal quantities, q^0 (and not q^1), are observed, it will not be possible to correctly estimate the discrete choice CV model by itself unless one can assume that there is no income effect (so that $q^1 = q^0$). But if, for exposition, it is temporarily assumed that

Table I

Descriptive Statistics for the Variables
(n = 3366)

Acronym	Description	Mean	Std. dev.
Y	median household income for respondent's 5-digit zip code (in $10,000)[a] (1980 Census scaled to reflect 1987 income; factor=1.699)	3.1725	0.9995
M	current trip market expenditures, assumed to be average for all trips (in $10,000)	0.002915	0.002573
T	annual lump sum tax proposed in CV scenario (in $10,000)	0.05602	0.04579
q	reported total number of salt water fishing trips to sites in Texas over the last year	17.40	16.12
I	indicator variable indicating that respondent would choose to keep fishing, despite tax T	0.8066	0.3950
DBAY	dummy variable for bays 1, 2, 4, and 5 (Sabine-Neches, Trinity-San Jacinto, San Antonio-Espiritu Santo, and Mission-Aransas)	0.4272	0.4948
PVIET	proportion of population in respondent's 5-digit zip code claiming Vietnamese ancestry	0.002497	0.006217
PFFF	proportion of population in respondent's 5-digit zip code employed in farming, forestry or fishing occupation	0.01277	0.01813

[a] Dollar-denominated quantities are expressed in $10,000 units throughout the study, so that squared income and squared net income do not become too large, resulting in extremely small probit coefficient estimates which thwart the optimization algorithm.

this is tenable, one can begin by considering how the model *might* be estimated.[11]

In anticipation of the joint models to be specified subsequently, it will be convenient to model the discrete choice elicited by the CV question using conventional maximum likelihood *probit* (rather than logit) techniques, where the underlying distribution of the implicit dependent variable, the true utility difference, is presumed to be Normal. Since ΔU in equation (7) can at best be only an approximation, assume that for the i^{th} observation, $\Delta U_i = \Delta U_i^* + \epsilon_i$, where ϵ_i is a random error term distributed $N(0, \sigma^2)$. ΔU_i^*, the linear-in-parameters systematic portion of the utility difference on the right hand side of equation (7) will be denoted in what follows as $x_i'\beta$ (to simplify the notation).

In conventional probit models, ΔU_i is unobserved, but if ΔU_i is "large" (i.e. $\Delta U_i > 0$), one observes an indicator variable, I_i (the "yes/no" response), taking on a value of one. Otherwise, this indicator takes the value zero. In constructing the likelihood function for this discrete response variable, the following familiar algebra is required:

$$(12) \qquad \Pr(I_i = 1) = \Pr (\Delta U_i > 0) = \Pr (\epsilon_i > - x_i'\beta).$$

Since ϵ_i has standard error σ, dividing through by σ will create a standard normal random variable, Z, with cumulative density function Φ.

$$(13) \qquad \begin{aligned} \Pr(\epsilon_i > - x_i'\beta) &= \Pr (Z > - x_i'\beta/\sigma) \\ &= \Pr (Z < x_i'\beta/\sigma) \\ &= \Phi (x_i'\beta/\sigma), \end{aligned}$$

by the symmetry of the standard normal distribution. The vector β can only be identified up to a scale factor, since it only ever appears in ratio to σ.

(However, this is quite acceptable, because the solutions to the consumer's utility maximization problem are invariant to monotonic transformations of the utility function.) The probability of observing $I_i = 0$ is just the complement of $Pr(I_i = 1)$, namely $1 - \Phi(x_i'\beta/\sigma)$, so the log-likelihood function for n observations will be:

$$(14) \quad \log L = \Sigma_i \; I_i \log [\Phi(x_i'\beta/\sigma)] + (1 - I_i) \log \{ 1 - [\Phi(x_i'\beta/\sigma)] \}$$

If q^1 could be observed, or if it could be legitimately assumed that income effects are zero, this separate discrete choice model could readily be estimated by any number of maximum likelihood routines in packaged statistical programs (such as SAS or SHAZAM) because the "index" $x_i'\beta$ is linear-in-parameters. For compatibility with what follows, however, this application requires a general MLE probit algorithm (in this case, using the GQOPT nonlinear function optimization package).

Unconstrained Parameter Estimates, Separate Demand Model

The demand function corresponding to a simple quadratic representation for the direct utility function is given by the formula in equation (9). The normalization $\beta_2 = 1$ results in equation (11). In what follows, the right hand side of equation (11) is denoted as $f(x_i,\beta)$, where the arguments imply a non-linear-in-parameters form.

In estimating this model separately, one might assume that $q_i = f(x_i,\beta) + \eta_i$, where η_i is normally distributed with mean zero and variance υ^2. The assumption that an additive $N(0, \upsilon^2)$ error is appended to (11) to generate the stochastic specification suggests that nonlinear least squares (by maximum likelihood) is the appropriate estimation method.

The log-likelihood function associated with the demand model is therefore:

(15) $\log L = -(n/2)\log(2\pi) - n \log \upsilon - (1/2)\ \Sigma_i\{\ [q_i - f(x_i,\beta)]/\upsilon\}^2$

Again, there exist packaged computational routines to estimate such nonlinear models, but this application requires a general function optimization program to allow for subsequent elaborations of this model.

Constrained Parameter Estimates, Independent Errors

The models described above would yield separate parameter estimates, reflecting the differences in the quadratic utility functions implied by each of the models taken *separately*.[12] However, to impose the requirement that the two decisions (one real and one hypothetical) reflect the *identical* underlying quadratic utility function, the models must be estimated simultaneously.

If one ignores the potential for unobserved heterogeneity to affect both decisions systematically, then the errors associated with each of the two models can be treated as independent. It is then simple to combine the two specifications. The joint probability of observing two independent events is simply the product of their individual probabilities. For each individual, then, one takes the product of the probability of his CV discrete choice and multiplies it by the probability density associated with his observed level of demand. This is the same as summing the two separate log-likelihood functions, but with the added feature that the corresponding β_j coefficients in each model are constrained to be the same. The log-likelihood function for this simplest joint model is therefore:

(16) $\log L = -(n/2)\log(2\pi) - n \log \upsilon - (1/2)\ \Sigma_i\ \{\ [q_i - f(x_i,\beta)]/\upsilon\}^2$

$\Sigma_i\ \{\ I_i \log\ [\Phi\ (x_i'\beta/\sigma)] + (1 - I_i) \log\ \{\ 1 - [\Phi\ (x_i'\beta/\sigma)]\ \}\ \}.$

Now, fortunately, it is possible to substitute the *formulas* for the fitted values of q^1 in place of the actual data on q^0 among the explanatory variables $x_i'\beta$ in the discrete choice portion of this log-likelihood.[13] It is not possible to maximize such a likelihood function (with parameter constraints) using any of the conventional prepackaged software programs, so estimation at this stage must use a general nonlinear function optimization program such as GQOPT.[14]

Constrained Parameter Estimates, Correlated Errors

Realistically, unobservable factors which affect respondents' answers to the CV discrete choice question are simultaneously likely to affect their actual number of fishing days demanded.[15] In the preferred model, in order to accommodate the influence of unmeasured variables, one can allow for a correlation, ρ, between the ϵ_i error terms in the discrete choice model and the η_i error terms in the demand model.[16] Assume that these errors have a bivariate normal distribution with parameters $(0, 0, \sigma^2, \upsilon^2, \rho)$.

In empirical discrete/continuous choice models, it is frequently more convenient *not* to work directly with the joint distribution of the errors. Instead, one can take advantage of the fact that the joint density can be represented equivalently as the product of a conditional density and a marginal density. In order to derive the model with nonzero ρ, one can exploit the fact that for a pair of standardized normal random variables, say W_1 and W_2, the conditional distribution of W_2, given $W_1 = w_1$, is univariate Normal with mean $(\rho\ w_1)$ and variance $(1 - \rho^2)$.

When allowing for nonzero values of ρ, then, the term $\Phi(x_i'\beta/\sigma)$ in the discrete-choice portion of equation (16) will be replaced by:

$$\Phi\ \{\ [(x_i'\beta/\sigma) + \rho\ Z_i]\ /\ (1 - \rho^2)^{1/2}\ \} \tag{17}$$

where

$$Z_i = [q_i - f(x_i,\beta)]/\upsilon, \tag{18}$$

the standardized fitted error in the demand function, evaluated at the current parameter values. (Changes in the sign reflect exploitation of the symmetry of the standard normal density function.) Clearly, if $\rho = 0$, this model collapses to the model with independent errors described in the previous section.

Alternative CV Interpretation: Price Change

The estimation of the model will be different under the potential alternative interpretation of the CV question acknowledged in the introduction. Respondents might infer that the CV question asks whether the they will continue to consume access days if the price of access is increased from M to $(M + T/q^0)$. As suggested above, a "yes" response implies that the respondent's optimal consumption of access days under the hypothesized scenario is positive. A "no" would mean that optimal consumption would actually be negative, but zero days are the fewest which can be consumed.

The "yes/no" response thus provides censored information regarding the magnitude of optimal quantity demanded. Unlike conventional probit models, where the location of the distribution is unknown (and therefore set arbitrarily to zero), the "threshold" in this case is exactly zero days. As above, $f(x_i,\beta)$ will be adopted as the generic representation for the Marshallian demand function corresponding to the quadratic utility model, where the variables x_i include income and the "price" of a day of access. As in Section 4, υ can be used as the same (constant) standard error of the conditional distribution of quantities demanded. The magnitude of υ can be inferred from observed consumption under current prices, so the dispersion of the unobservable dependent variable in the CV model is "known" (in contrast to the conventional probit situation).

Providing, then, that it is reasonable to assume that real and hypothetical behavior derive from the identical set of underlying preferences, the discrete responses to the CV question can be used to supplement the estimation of the underlying demand parameters. Specifically, the expression $(x_i'\beta/\sigma)$ in equations (16) and (17) will be replaced by $f(x_i^*,\beta)/\upsilon$, where the elements of x_i^* reflect current actual income, but price M is replaced by the hypothesized $(M+T/q^0)$.

One difference under this interpretation of the CV question is that this specification no longer allows identification of the individual utility parameters (β_1 through β_5, up to the scale factor, σ, of the unobservable dispersion in the latent variable driving the CV response). Only the demand parameters, β_1^*, β_3^*, β_4^*, and β_5^* and υ can be identified. Fortunately, the utility function is invariant to the scale of the parameters and arbitrarily setting $\beta_2 = 1$ will result in exactly the same implications in terms of optimizing behavior.

5. CLASSES OF RESULTS AND GENERAL DISCUSSION

The central empirical results in this study are the estimates of the parameters of the assumed underlying quadratic direct utility function. All of the economically interesting empirical measurements in this paper are derived from this utility function. Throughout, the empirical utility function should exhibit properties which are consistent with economic intuition about plausible shapes for these functions.

Utility Derivatives

First, the derivatives of the underlying direct utility function (equation 3) are:

(19) $$\partial U/\partial z = \beta_1 + \beta_3 z + \beta_4 q \qquad \partial^2 U/\partial z^2 = \beta_3$$
$$\partial U/\partial q = \beta_2 + \beta_4 z + \beta_5 q \qquad \partial^2 U/\partial q^2 = \beta_5$$
$$\partial^2 U/\partial z \partial q = \beta_4$$

The marginal utilities of the composite good z and of fishing days q should be positive, but the specific values of these marginal utilities will depend on the local values of z and q. Whether or not each marginal utility is increasing or decreasing will be revealed by the signs of β_3 and β_5. Eventually, one would expect diminishing marginal utility to set in. Within the estimating sample, however, it is entirely plausible that constant or even increasing marginal utility might obtain.

Global Optima or Saddle Points

If both β_3 and β_5 are negative, the fitted utility function will be globally concave, and a globally optimal combination of z and q, (z^*,q^*) will be implied. The consumer's optimization problem will be constrained unless the bliss point is attainable inside the budget constraint. The formulas for the bliss point will be strictly in terms of the estimated coefficients:

(20) $$q^* = [-\beta_2 + (\beta_1 \beta_4/\beta_3)] \;/\; [\beta_5 - (\beta_4^2/\beta_3)]$$
$$z^* = (-\beta_1 - \beta_4 q^*)/\beta_3$$

For any set of fitted utility parameters implying a strictly concave preference mapping, it will be helpful to check the location of the bliss point relative to the range of the actual data on z and q.

Admissible fitted quadratic utility functions are not necessarily strictly concave, however. The point (q^*, z^*) may correspond to a saddle point of the utility function. But only quasi-convexity is required. To assess compliance with this regularity condition, one can easily examine the configuration of the fitted utility function's indifference curves.

Indifference Curves

It is straightforward to compute an indifference curve through any arbitrarily chosen bundle (z', q'). First, determine the level of utility this bundle represents:

$$U' = \beta_1 z' + \beta_2 q' + \beta_3 z'^2/2 + \beta_4 z'q' + \beta_5 q'^2/2 \tag{21}$$

Having found the value of U', set up a quadratic formula for z:

$$(\beta_3/2)z^2 + (\beta_1 + \beta_4 q)z + [\beta_2 q + (\beta_5/2)q^2 - U'] = 0 \tag{22}$$

Solving this quadratic formula yields z as a function of q and the (constant) U'. For an arbitrarily chosen range of values for q, then, the formula can be used to generate the corresponding values of z which lie along the same indifference curve as does (z', q'). These values can then be plotted for a visual display of the preference configuration (through its level curves). For example, it is particularly interesting to examine the indifference curve through $(\mu_Y, 0)$, where μ_Y is the sample mean income (equal to μ_z if q is zero). Also informative will be the budget constraint through this point corresponding to the sample mean market expenditures, μ_M. Finally, the optimal quantity q under current income μ_Y and "price" μ_M, can be calculated and the indifference curve through this point can be included in the diagram. These two indifference curves and the budget constraint convey the intuition behind using the fitted models to derive the "break-even" value of access, T* (i.e. the equivalent variation for a complete loss of access).

Demand Functions

Marshallian demand functions are also produced by these models. Once a demand function has been calibrated, certain of its properties are usually of interest, such as the implied price and income derivatives:

$$\partial q/\partial p = [\ -(2\beta_4 M-\beta_3 M^2-\beta_5)(\beta_1+\beta_3 Y)\ -\ 2(\beta_2+\beta_4 Y-\beta_1 M-\beta_3 MY)(\beta_4-\beta_3 M)\ ]\ /\ [2\beta_4 M-\beta_3 M^2-\beta_5]^2 \tag{23}$$

$$\partial q/\partial Y = [\beta_4-\beta_3 M]\ /\ [2\beta_4 M-\beta_3 M^2-\beta_5]^2$$

where one could alternatively set $\beta_2 = 1$ and describe these quantities in terms of the β_j*. Price and income elasticities at each point (or at the means of the data) will clearly depend upon q, p, and Y for each observation.

Policy makers may also be interested in estimates of the "choke price." Depending upon whether one is concerned with the total price of access, (M + p), or just the potential increases in price, p, which might be instituted through per-day access charges, the procedure will be similar. In the first case, given income, one simply sets q = 0 in equation (9) and solves the resulting quadratic formula for (M + p). If the current level of M is known, the "choke" level of a potential per-day access charge can readily be determined.

Equivalent Variation for Complete Loss of Access

One motivation for this paper concerns estimation of the total social value of recreational access to this fishery. One measure of value is the tax, T* (the equivalent variation) which would make these anglers just indifferent between paying the tax and continuing to fish, or not paying the tax and forgoing their fishing opportunities. The estimated utility parameters for the fitted model can be used to determine *for each respondent* the tax level T* which would make that individual just indifferent between the two options offered. This T* is simply the value of T, when combined with the data on Y, M, and q^1, for this individual, makes the fitted utility difference expression, ΔU in equation (7), equal to zero. One might be tempted simply to solve (7) for T which yields the quadratic expression:

$$(24) \qquad 0 = a\,T^2 + b\,T + c = (\beta_3/2)\,T^2 - [\beta_1+\beta_4 q+\beta_3(Y-Mq)]\,T + \{\,(\beta_3/2)[(Y-Mq)^2-Y^2] + (\beta_5/2)q^2 + [\beta_2-\beta_1 M-\beta_4(Y-Mq)]\,q\,\}$$

However, it is not *measured* q^0, but *post-tax* optimal quantity, q^1, which relevant to this formula. Wherever quantity appears in equation (7), then, one must substitute the fitted value of q^1, which is given by:

$$(25) \qquad q^1 = [\,\beta_2 + \beta_4(Y - T) - \beta_1\,M - \beta_3 M(Y - T)]\,/\,[\,2\,\beta_4 M - \beta_3 M^2 - \beta_5\,].$$

Since this expression also involves T, the coefficients of the resulting quadratic form for T* will be somewhat more complicated. They are straightforward, however. Letting $D = (2\beta_4 M - \beta_3 M^2 - \beta_5)$, let $R = (\beta_2+\beta_4 Y-\beta_1 M-\beta_3 MY)/D$ and $S = (\beta_4-\beta_3 M)/D$, the quadratic form for T* can be written:

$$\begin{aligned}(26) \quad 0 = {} & [\,(\beta_3/2)(MS-1)^2 - \beta_4 S(MS-1) + (\beta_5/2)S^2\,]\,T^{*2} \\ & + [\,\beta_1(MS-1) - \beta_2 S + \beta_3(Y-MR)(MS-1) + \beta_4\{R(MS-1) - (Y-MR)S\} - \beta_5 RS\,]\,T^* \\ & + [\,-\beta_1 MR + \beta_2 R + (\beta_3/2)\{(Y-MR)^2 - Y^2\} + \beta_4(Y-MR)R + (\beta_5/2)R^2,\end{aligned}$$

Where Y is income and M denotes associated market expenses, as before. Solving this quadratic formula for T* for each respondent yields a "best guess" regarding the equivalent variation associated with access to the fishery enjoyed by each person in the estimating sample. If the sample was truly representative of the population of resource users, then an estimate of the total social value accruing to current users could be obtained by scaling the aggregate sample value up to the population.

Equivalent Variation for Partial Loss of Access

Completely depriving everyone of access to the resource is an extremely drastic proposition. Fortunately, the models described in this paper are also well-suited to evaluating the social costs of modest decreases in access. This section generalizes by considering how to use the fitted models to assess the social costs of limiting access to a proportion α ($0 < \alpha < 1$) of current (fitted) access levels. Using $q*$ to denote current access days, utility *after* the restriction will be:

$$U_\alpha = \beta_1 (Y - \alpha M q*) + \beta_2 \alpha q* + \beta_3 (Y - \alpha M q*)^2/2 + \beta_4 (Y - \alpha M q*)\alpha q* + \beta_5 \alpha q*^2/2, \tag{27}$$

where

$$q^* = [\beta_2 + \beta_4 Y - \beta_1 M - \beta_3 MY] / [2 \beta_4 M - \beta_3 M^2 - \beta_5] \tag{28}$$

evaluated at the optimized parameter values for the basic model.

If the intervention was not access restriction, but instead an income reduction, one could ask just what "tax," T' (equivalent variation), would result in the consumer's freely optimized utility level ($U_{T'}$) being the same as U_α, where

$$U_{T'} = \beta_1 (Y-Mq'-T') + \beta_2 q' + \beta_3 (Y-Mq'-T')^2/2 + \beta_4 (Y-Mq'-T')q' + \beta_5 q'^2/2. \tag{29}$$

and

$$q' = [\beta_2 + \beta_4(Y-T') - \beta_1 M - \beta_3 M(Y-T')] / [2 \beta_4 M - \beta_3 M^2 - \beta_5]. \tag{30}$$

Substituting the expressions for q* and q', and then setting $U_\alpha = U_{T'}$, it is tedious but straightforward to solve for the value of T' which satisfies the equality. The solution is again a quadratic formula:

$$\begin{aligned}(31)\quad 0 = {} & [\ (\beta_3/2)(MS-1)^2 - \beta_4 S(MS-1) + (\beta_5/2)S^2\ ]\ T'^2 \\ & + [\ \beta_1(MS-1) - \beta_2 S + \beta_3(Y-MR)(MS-1) + \beta_4\{R(MS-1) - (Y-MR)S\} - \beta_5 RS\ ]\ T' \\ & + [\ -\beta_1(1-\alpha)MR + \beta_2(1-\alpha)R + (\beta_3/2)\{(Y-MR)^2 - (Y-\alpha MR)^2\} \\ & + \beta_4\{(Y-MR)R - (y-\alpha MR)(\alpha R)\} + (\beta_5/2)(1-\alpha^2)R^2\ ],\end{aligned}$$

where R and S are as defined in the case of complete loss of access. Only the constant term in the quadratic expression is affected by the introduction of the proportion α. When $\alpha=0$, the formula reduces to that given above in (26). Since the coefficients in this quadratic formula are simply data and estimated parameters, the fitted value of T' can readily be computed for any arbitrarily specified value of α.

Compensating Variation for Complete Loss of Access

The equivalent variations give the most that a respondent would be willing to pay before he would elect to discontinue consuming the resource. An alternative measure of value asks what sum of money would have to be given to a respondent who has been denied access in order to leave him equally well off as before the intervention. Algebraically, this is the amount of money, C, which would make current utility, U (under income Y and travel costs M), just equal to utility U_C (with zero access days but supplemented income Y+C).

The compensating variation C can be determined by finding the root of the quadratic formula:

$$\begin{aligned}(34)\quad 0 = {} & -(\beta_3/2)\ C^2 - (\beta_1 + \beta_3 Y)\ C \\ & -\beta_1 Mq + \beta_2 q + (\beta_3/2)[\ (Y-Mq)^2 - Y^2] + \beta_4(Y-Mq)q + (\beta_5/2)q^2.\end{aligned}$$

In this case, q is merely *current* fitted quantity, q^0, rather than the endogenously determined post-tax quantity employed in the formulas in the previous section.

6. SPECIFIC EMPIRICAL ESTIMATES

The Basic Model; Equivalent and Compensating Variations

The basic model restricts the quadratic direct utility parameters and the corresponding parameters in the Marshallian demand function for fishing days to be identical. The model assumes equal reliability of the two types of information (CV and actual market demand), and allows the post-tax quantity demanded in the discrete choice model to be determined endogenously according to the same demand function. The model also allows for correlated errors in the two decisions. The *first* column of Table IIa gives these results (subsequent columns will be discussed in a later section).

The utility function implied by these parameter estimates is globally concave, with a slightly positively sloped principal axis of the ellipses which form its level curves (the indifference curves). Of course, the quadratic form is merely a local approximation to the true utility function. Nevertheless, if the entire surface of the true utility function was quadratic, the apparent "bliss point" of that function would be located at 28.4 fishing days and $289,823 in median zip code income (compared to sample means of 17.4 fishing days (highly skewed) and $31,725 income. At the means of the data, the two marginal utilities are positive. The implied price elasticity of demand at the means is -0.074 and the income elasticity is 0.078. To establish a visual benchmark for this basic model, for an individual with mean income and travel costs, an indifference curve for the empirical quadratic utility function, the budget constraint through $(\mu_Y,0)$, and the fitted maximum attainable indifference curve are shown in Figure 1.

Table IIa

Auto-Validated Joint Discrete Choice/Demand Model

Parameter	Point Est. (Asymp. t-ratio) reliability=1.0[a]	Point Est. (Asymp. t-ratio) reliability=0.5	Point Est. (Asymp. t-ratio) reliability=0.1
β_1 (z)	3.309 (8.237)	4.416 (8.200)	7.840 (6.385)
β_2 (q)	0.1192 (19.55)	0.1253 (18.47)	0.1399 (12.64)
β_3 ($z^2/2$)	-0.1167 (-1.836)	-0.3404 (-2.353)	-1.036 (-2.986)
β_4 (zq)	0.002579 (2.006)	0.001421 (0.9798)	-0.001093 (-0.6008)
β_5 ($q^2/2$)	-0.006837 (-22.80)	-0.006891 (-21.48)	-0.007060 (-13.47)
υ	16.01 (81.98)	16.00 (94.69)	15.98 (110.5)
ρ	0.2315 (9.086)	0.2316 (7.476)	0.2324 (4.030)
max Log L	-15708.17	-19878.36	-25938.13

[a] "Reliability" is the size of the weight on the hypothetical CV information relative to the weight on the observed demand behavior.

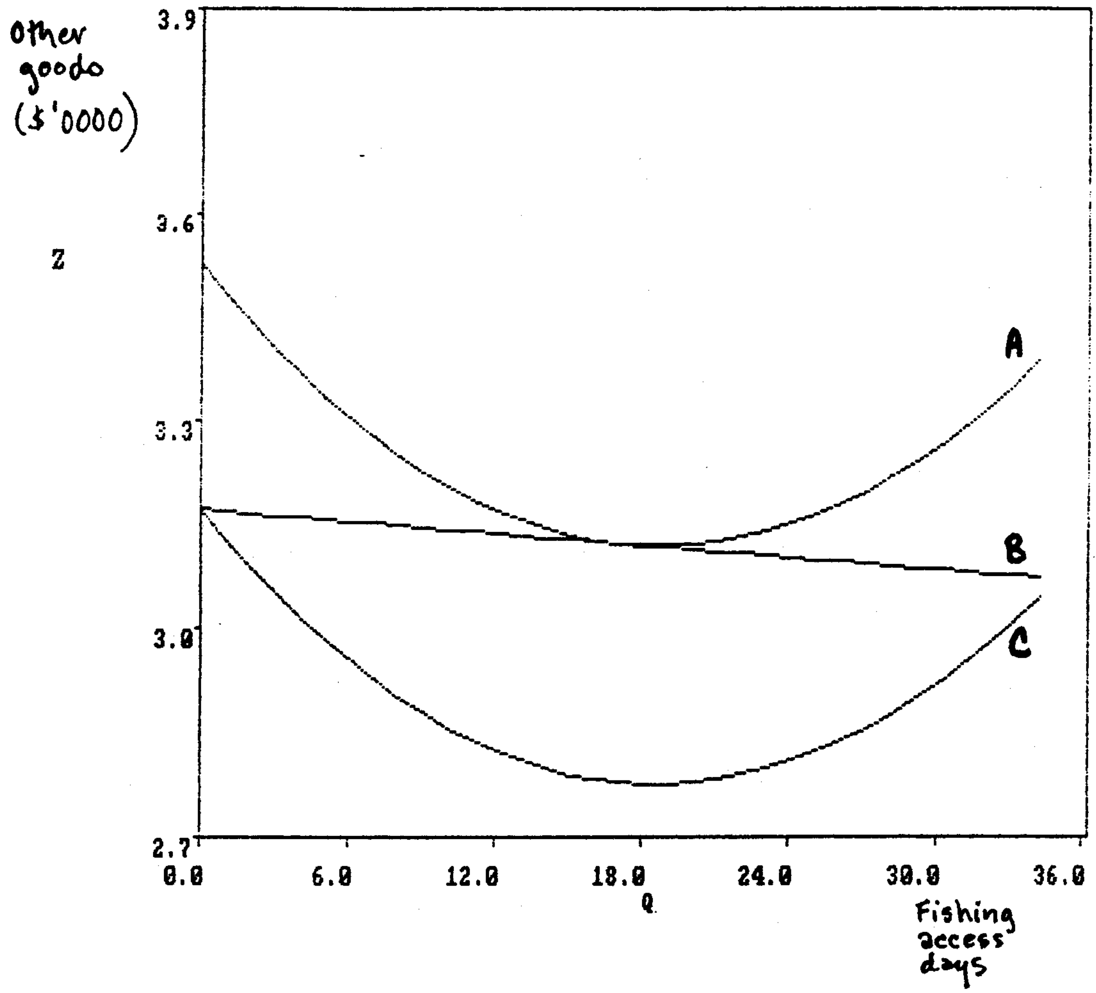

Figure 1 - For respondent with mean income and travel costs and preferences estimated using "basic model": indifference curve at current optimum (A); current budget constraint (B); indifference curve associated with exclusion from resource (C). Note: z = other goods ($10,000), q = access days.

Using the basic model, it is possible to compute fitted values for T* for each respondent (the equivalent variation, or "tax," which would make the respondent just as badly off as a complete forfeiture of fishing access). Across the 3366 respondents in this sample, descriptive statistics for the fitted values of T* appear in the first row of Table IIb.[17] Over the entire sample, the average equivalent variation for a complete loss of access is $3451.

However, since a complete abolition of resource access is an unlikely event, it is sensible also to consider the model's implications regarding successively smaller restrictions on days of access. If α denotes the proportion of current consumption to which each individual's access days are restricted, subsequent rows of Table IIb give descriptive statistics for the fitted measures of equivalent variation for each change.[18] For an across-the-board 10% reduction in fishing days, for example, the average utility loss by these respondents would be only $35.

The final row of Table IIb provides, for comparison, the corresponding compensating variation for a complete loss of access ($\alpha = 0$). As is typical, the compensating variation for the loss is larger than the equivalent variation for the same loss. Here, however, the difference is largely an artifact of the globally concave quadratic form of the fitted utility function.

Extension: CV Question Interpreted as Price Change

The parameter estimates for the utility function under this fundamentally different interpretation of the CV question appear in Table III. Only the "demand" form of the parameters is estimable, and it is not surprising that the point estimates differ systematically from their counterparts in the next section.

Table IIb

Fitted Equivalent and Compensating Variation Estimates; Basic Model (Table IIa) (DESCRIPTIVE STATISTICS, N = 3366)

Valuation Measure:	mean	std. dev.	max	min
Equivalent Variation				
$\alpha = 0.0$[a]	$ 3451	$ 509	$ 5132	$ 1857
$\alpha = 0.1$	2799	414	4166	1505
$\alpha = 0.2$	2214	328	3298	1190
$\alpha = 0.3$	1697	251	2529	912
$\alpha = 0.4$	1248	185	1861	670
$\alpha = 0.5$	867	129	1294	465
$\alpha = 0.6$	555	82	829	298
$\alpha = 0.7$	313	46	467	168
$\alpha = 0.8$	139	21	207	75
$\alpha = 0.9$	35	5	52	19
Compensating Variation				
$\alpha = 0.0$	$ 3560	$ 535	$ 5361	$ 1899

[a] For access days restricted to the fraction α of fitted current access days.

Table III

Model with CV Question Interpreted as Price Change

Parameter	Point Estimate (asymp. t-ratio)
β_1^* (z)	19.80 (5.366)
β_2^* (q)	1.000 -
β_3^* ($z^2/2$)	-2.613 (-2.573)
β_4^* (zq)	0.03155 (1.726)
β_5^* ($q^2/2$)	-0.06163 (-18.23)
υ	16.18 (86.75)
ρ	0.08754 (3.080)
Max. LogL	-15708.12

For this version of the joint model, the marginal utilities at the means of the data are positive; the price elasticity of demand for access days is -0.035; the income elasticity is 0.11. The implied bliss point is 20.2 access days and $78212 in median household income.

While the fitted utility function under this interpretation is completely plausible from a theoretical standpoint, the implications of this model are quite a bit different from the "lump-sum tax" CV interpretation. The sample mean of the fitted equivalent variations for a complete loss of resource access, according to these preferences, is markedly higher, at $7386 (with standard deviation $2244). Clearly, subsequent surveys will have to be very careful in conveying to respondents exactly what type of scenario is intended, since the interpretation of the question can make almost an order of magnitude difference in the results.

Extension: Differing Reliability for Real versus Contingent Data

The basic model reflects the presumption that the decisions which respondents claim they would make under the hypothetical scenario proposed in the CV question deserve to be treated as *equally credible* when compared to their actual market behavior regarding number of fishing days demanded. This need not be the case. Contingent valuation data have regularly been criticized as potentially unreliable because they represent "hypothetical answers to hypothetical questions." With auto-validation, however, the empirical researcher is free to adjust the relative weights on each type of information to reflect prior opinions regarding the reliability of the contingent valuation responses.

Under equal credibility, the terms in the log-likelihood function corresponding to each type of information each receive an implicit unit weight. However, if intuition suggests instead that the available CV information is only *half* as reliable as real market information, one might weight the CV term in the log-likelihood function by 2/3 and the demand term by 4/3, so that the weights still sum to two. (This will be designated as a "reliability" factor of .5 for the CV information.) The point estimates of the parameters are certainly sensitive to these weights. The heavier the weight on the "real" data, the more the point estimates of the parameters will look like the unrestricted estimates for the demand equation by itself. In contrast, the greater the weight on the CV data, the more the point estimates will look like the unrestricted estimates for the discrete choice utility parameters estimated separately. The ultimate extremes of course, will occur when a zero weight is placed on one term and a weight of two is applied to the other. (The likelihood function in that case will be simply twice the function for either model estimated separately--the same parameters will result.) However, it is not possible to allocate zero weight to the demand portion of the model because it would no longer be possible to endogenize the post-tax optimal quantities in the CV part of the model. It appears *not* to be possible both to endogenize quantity and to estimate the utility parameters with only the CV data.[19]

Recognizing the effects of varying the weights suggests a useful conceptual experiment. Given the maintained hypothesis of a quadratic utility function (which is admittedly only an approximation), one can ask just how small the weight on the CV information would have to become before LR tests could just fail to reject the null hypothesis of parameter equivalence for the two decisions. This involves a line search across possible values of the weight on the CV term in the log-likelihood function.[20]

Using likelihood ratio test statistics to determine whether the utility parameters implied by the CV data "alone" are consistent with those estimated jointly using both CV and demand data requires a "restricted" and an "unrestricted" specification. The restricted specifications for reliability equal to 1.0 (the basic model), 0.5 and 0.1 appear in Table IIa. For the "unrestricted" models, *some* demand information is unfortunately necessary to correctly estimate utility function parameters from CV model (unless one assumes that income effects are zero, which would confound the desired hypothesis). The best compromise seems to be to allow the discrete choice CV model to imply values for β_1, β_2, β_3, β_4, and β_5. Likewise, the observed demand decisions will imply values for β_1*, β_3*, β_4*, and β_5* (but not β_2). For the *endogenously* determined quantity in the CV discrete choice model, however, one must "borrow" the parameters from the demand model: β_1*, β_3*, β_4*, and β_5*. Table IV displays sets of results for these "unrestricted" models, corresponding to each column in Table IIa.

First of all, for *equal* weights applied to the two types of information (CV and demand data) it is useful to compare the utility functions implied by the CV responses with and without parameter restrictions. For the "average" respondent, Figure 2a shows the indifference curve associated with zero access days for the basic (restricted) model (curve D) as well as indifference curves corresponding to the utility parameters generated by (i.) the CV portion of the unrestricted model (curve A) and (ii.) the demand portion of the unrestricted model (curve C). The greater curvature of the indifference curve for the unrestricted CV parameters implies that T* (the "break-even" flat tax, or equivalent variation) based primarily on the CV responses, will be substantially larger than T* based on observed market demand behavior. Interestingly, however, the basic auto-validated model yields equivalent variation estimates very similar to those from the unrestricted CV parameters.[21] Specifically, for the latter parameters, the sample mean fitted

Table IV

Model with Distinct CV Utility and Observed Demand Parameters
(Endogenous CV Demand Given by Demand Function Parameters)

Parameter	Point Est. (Asymp. t-ratio) reliability=1.0[a]	Point Est. (Asymp. t-ratio) reliability=0.5	Point Est. (Asymp. t-ratio) reliability=0.1
β_1 (z)	4.510 (4.435)	5.356 (5.814)	6.772 (3.826)
β_2 (q)	0.3223 (11.09)	0.3300 (10.64)	0.3378 (6.333)
β_3 ($z^2/2$)	0.4937 (1.680)	0.2011 (0.8696)	-0.3299 (-0.7179)
β_4 (zq)	0.01185 (3.076)	0.008439 (2.267)	0.003184 (0.6130)
β_5 ($q^2/2$)	-0.03078 (-9.473)	-0.03050 (-9.326)	-0.02971 (-5.406)
$\beta_1^* = \beta_1/\beta_2$	46.51 (8.894)	55.51 (10.54)	69.82 (10.70)
$\beta_3^* = \beta_3/\beta_2$	0.1819 (1.767)	-3.483 (-2.901)	-9.347 (-4.604)
$\beta_4^* = \beta_4/\beta_2$	0.02705 (1.821)	0.01079 (0.8029)	-0.01229 (-1.034)
$\beta_5^* = \beta_5/\beta_2$	-0.05448 (-20.81)	-0.05198 (-22.73)	-0.04853 (-24.21)
υ	15.98 (81.90)	15.98 (94.71)	15.97 (110.6)
ρ	0.2483 (9.659)	0.2478 (7.724)	0.2462 (4.002)
max Log L	-15634.21[b]	-19825.80[c]	-25920.08[d]

[a] "Reliability" is the size of the weight on the hypothetical CV information relative to the weight on the observed demand behavior.

[b] Compared to results in Table IIb, LR test for hypothesis of same β parameters for utility and demand functions is 147.92 (when the 5% critical value of the χ^2 test statistic is 9.49 and the 1% critical value is 13.28).

[c] LR test for same β parameters is 105.12; reject.

[d] LR test for same β parameters is 36.1; reject.

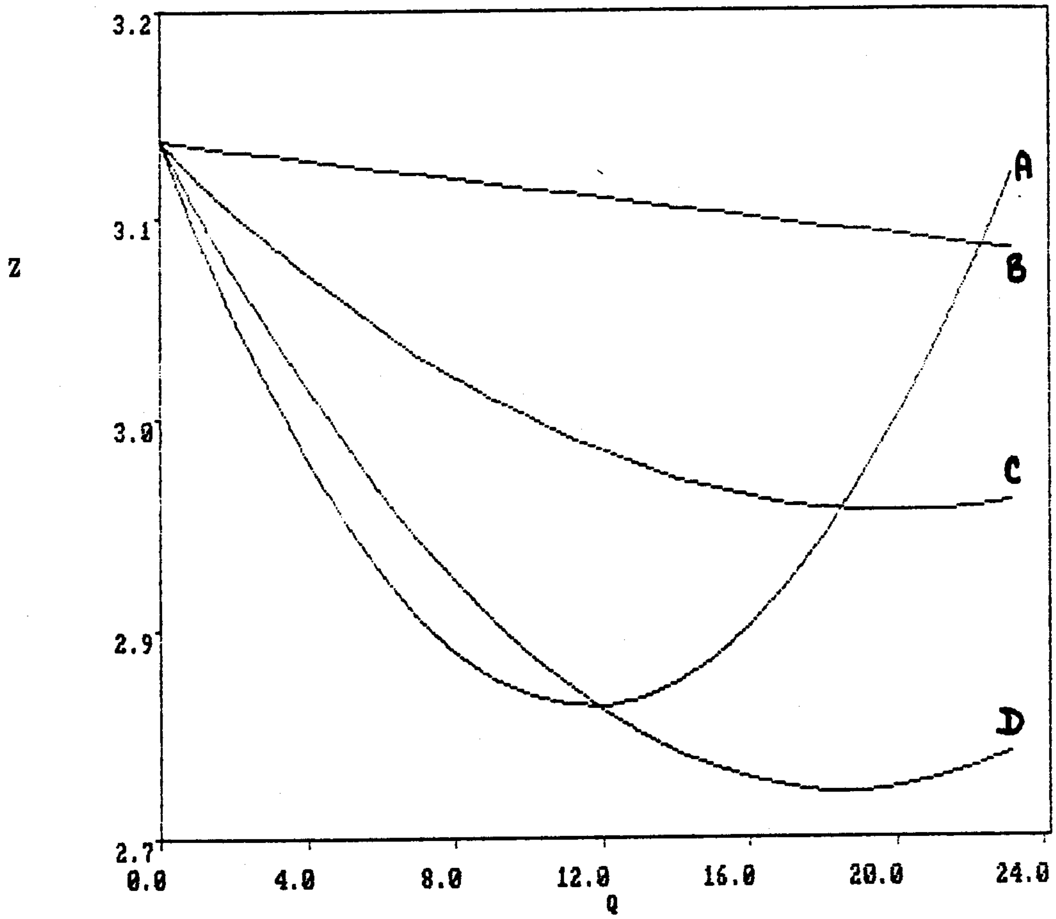

Figure 2a - For respondent with mean income and travel costs: current budget constraint (B); indifference curves associated with exclusion from resource based on different sets of utility parameters--basic model (D), contingent valuation portion of unrestricted model (A), and demand portion of unrestricted model (C).

equivalent variation is $3118 (standard deviation $283) versus $3451 for the restricted model. For the unrestricted demand parameters, the sample mean fitted equivalent variation is only $1762 (standard deviation $412).

The implied *inverse* demand functions corresponding to each of the three sets of preferences are shown in Figure 2b. Quantity demanded at $p = 0$ (i.e. at total access price $M + p = M$) is the fitted optimum for this consumer under current conditions. The unconstrained CV parameters imply a much higher choke price (over $542) and much less elastic demand than does the utility function estimated when CV responses and observed demand are constrained to reflect the same set of quadratic preferences (choke price about $409). Unrestricted actual demand behavior implies a lower choke price (about $219). The pure CV (i.e. hypothetical market) scenario *does* seem to invite respondents to overstate the strength of their demand for resource access, as one might suspect.

But these findings pertain to models wherein real and hypothetical responses are given equal weight. What about the degree of convergence in the estimates as the hypothetical CV responses are down-weighted? The footnotes to Table IV describe the results. The relative reliability of the CV information has been decreased to 0.1 and it is still possible to reject the hypothesis of common utility parameters. The restricted model constrains four parameters which are allowed to take on whatever values the data suggest in the unrestricted models. The difference in the restricted and unrestricted log-likelihood functions would have to fall to roughly 5 or 6 before the LR test would fail to reject the hypothesis of parameter equivalence in the CV and demand models. It is interesting, then, to observe that it would be quite a "stretch" to bring the implications of the hypothetical CV responses into line with the observed demand behavior.[22]

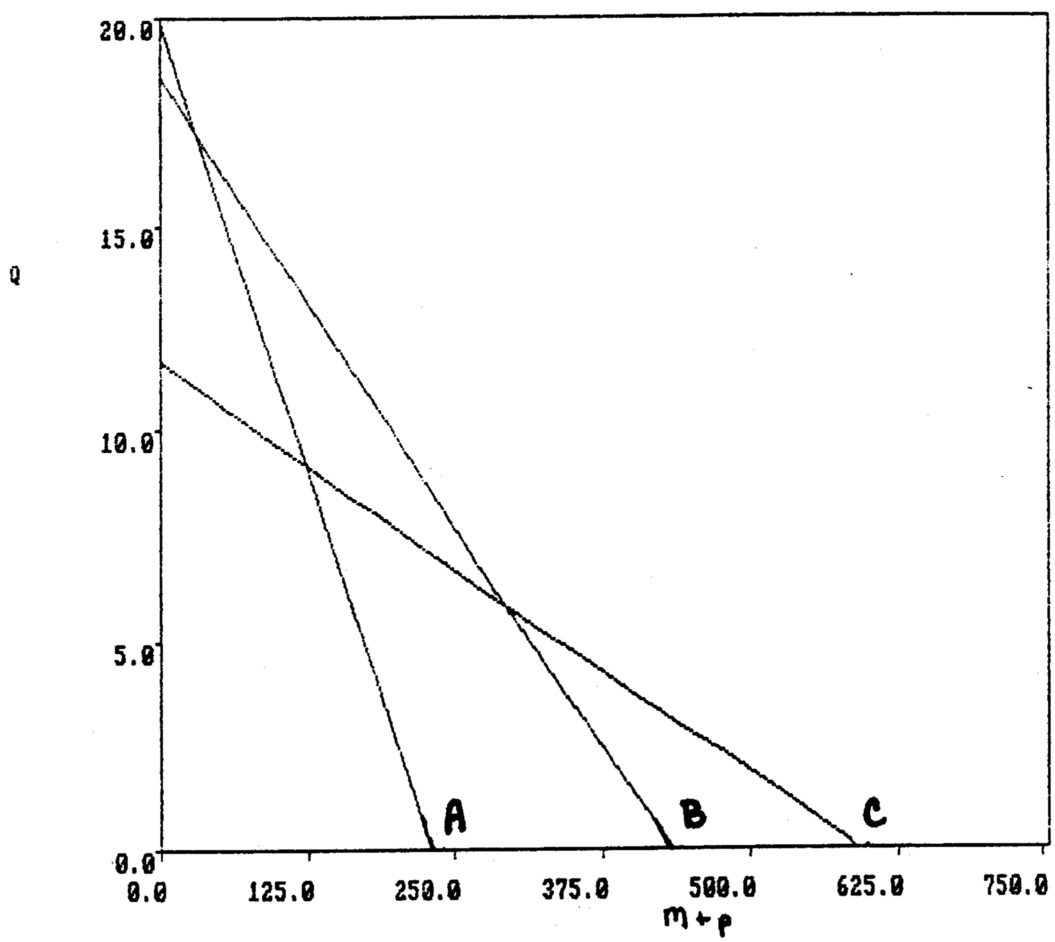

Figure 2b - For respondent with mean income and travel costs: demand curves implied by different sets of utility parameters--basic model (B), contingent valuation portion of unrestricted model (C), and demand portion of unrestricted model (A).

Extension: Estimates in the absence of travel cost data

In some applications, M may be measured accurately and may be relatively constant across fishing days, but in other cases, it may not. Sometimes, the researcher may be better off ignoring the questionable information on M, and using a simpler "Engel curve" model as opposed to a "demand function" (where equation numbers indicate revisions of the original specification):

$$(1') \qquad U(Y - T, q^1) > U(Y, 0).$$

$$(2') \qquad \Delta U = U(Y - T, q^1) - U(Y, 0) > 0.$$

If the data on M are excluded, $\dot{z}$ will be identically Y.

$$(6') \qquad \Delta U = U_T - U_N = \beta_1 [Y-T] + \beta_2 q^1 + \beta_3 [Y-T]^2/2 + \beta_4 [Y-T]q^1 + \beta_5 (q^1)^2/2 - \{ \beta_1 [Y] + \beta_3 [Y]^2/2 \}.$$

$$(7') \qquad \Delta U = \beta_1 \{[Y-T] - Y\} + \beta_2 q^1 + \beta_3 \{[Y-T]^2 - Y^2\}/2 + \beta_4 [Y-T]q^1 + \beta_5 (q^1)^2/2.$$

$$(8') \qquad L = \beta_1 z + \beta_2 q + \beta_3 z^2 + \beta_4 zq + \beta_5 q^2 + \lambda (Y - pq).$$

$$(9') \qquad q = [\beta_2 + \beta_4 Y - \beta_1 p - \beta_3 Y p] / [2 \beta_4 p - \beta_3 p^2 - \beta_5].$$

$$(11') \qquad q^0 = [1 + (\beta_4{*}) Y] / [- \beta_5{*}].$$

$$(23') \qquad \partial q/\partial p = [\beta_5(\beta_1-\beta_3 Y) - 2\beta_4(\beta_2+\beta_4 Y)] / [\beta_5]^2$$

$$\partial q/\partial Y = -\beta_4/\beta_5.$$

In order to appreciate the benefits of joint estimation with income data and numbers of trips but in the absence of travel costs as proxy data for prices, one can consider the estimates of the utility function parameters when the data on M in this sample are ignored. Table V displays these results. At the means of the data, these fitted parameters imply a utility function with positive marginal utility from other goods, but very slightly negative marginal utility from access days. This implies that the utility function in this case is not globally concave. The saddle point of the utility function is located at 12.25 access days and \$-47348. Nevertheless, the level curves are still convex to the origin. At the means of the data, the price elasticity of demand for access days is -0.125 and the income elasticity is 0.0682.

Figure 3 shows the effects on the fitted preference function of ignoring travel costs in the estimation phase. As benchmarks, this figure includes the "basic" indifference curve for a typical respondent (curve E) as well as the indifference curve based on the CV portion (curve A) and the *demand* portion (curve D) of the unrestricted model. Here, however, attention should be focused on the indifference curve for a model similar to the basic model except that the available data on travel costs are ignored (curve A). Even this very "thin" information about market demand pulls the parameter estimates a long way away from the unrestricted CV estimates depicted by curve A. Still, it is not clear in this application that the resulting (much smaller) equivalent variation estimates will be superior to those generated by the CV portion of the unrestricted model.

Extension: Accommodating Respondent Heterogeneity

All of the models described above have presumed that these respondents are homogeneous on all dimensions other than income, Y, proposed tax, T, number of fishing days, q, and typical individual market expenditures, M. It is a simple matter, however, to relax this assumption, allowing the parameters

Table V

Jointly Estimated Model Ignoring Travel Costs (i.e. M = 0; Only Engel Curves from Observed Demand Employed)

β_1 (z)	3.586
	(1.342)
β_2 (q)	0.1259
	(13.19)
β_3 $(z^2/2)$	0.7711
	(0.9538)
β_4 (zq)	0.005329
	(2.058)
β_5 $(q^2/2)$	-0.008213
	(-22.46)
υ	16.12
	(81.85)
ρ	0.2343
	(9.076)
log L	-15679.17

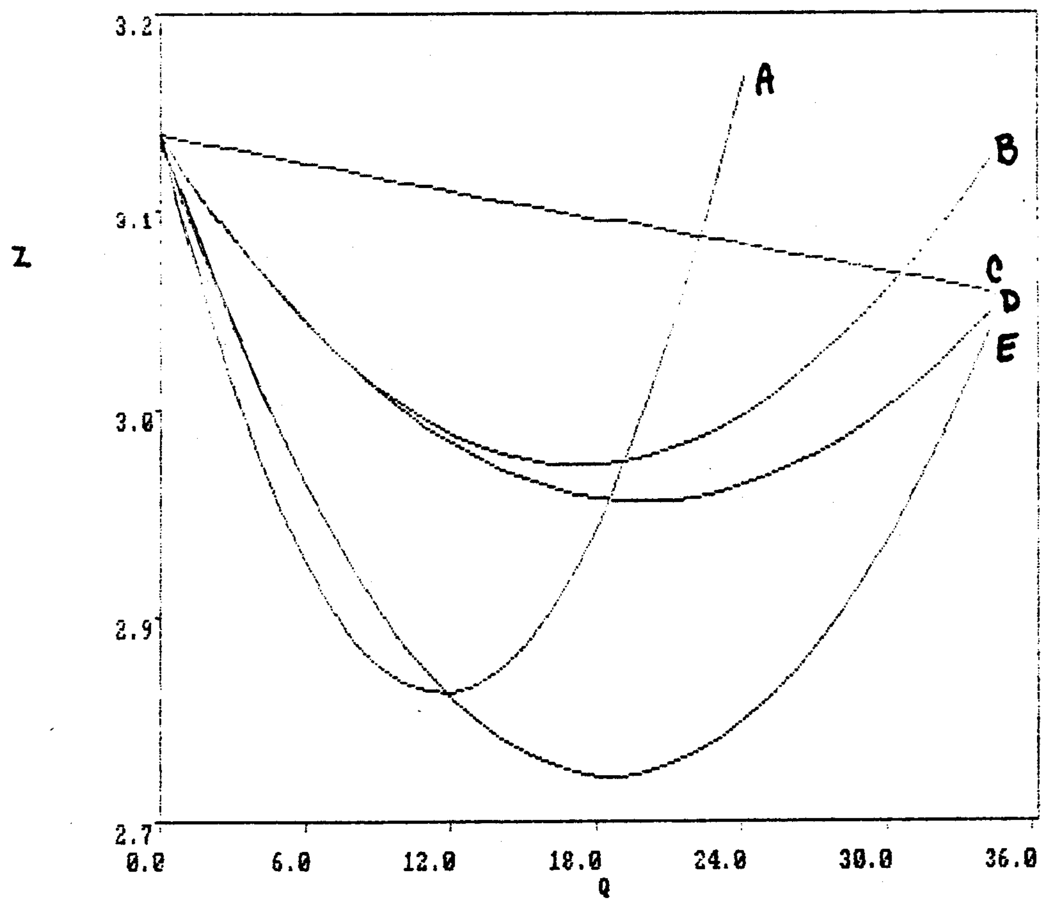

Figure 3 - For respondent with mean income and travel costs, effects of ignoring travel costs during estimation of utility parameters by modified basic model: actual budget constraint (C), indifference curve from basic model (E), indifference curve from the CV portion of the unrestricted model (A), indifference curve from demand portion of unrestricted model (D), and indifference curve from model estimated without travel cost (using only Engel curve information) (B).

of the quadratic utility function to vary systematically with the levels of other measured attributes, X, of either the respondent or the resource.

For example, there are eight major bay systems in the sample area. Other work with the full data set (Cameron, Clark, and Stoll, 1988) has suggested that perhaps one set of bays is systematically different from the rest. In the first example, consider a dummy variable X = DBAY taking on the value of 0 for bays 3, 6, 7, and 8 (Lavaca-Tres Palacios, Corpus Christi-Neuces, Upper and Lower Laguna Madre) and the value 1 for bays 1, 2, 4, and 5 (Sabine-Neches, Trinity-San Jacinto, San Antonio-Espiritu Santo, and Mission-Aransas). Allow this variable to shift the parameters of the quadratic utility function, so that β_j is replaced by $\beta_j + \gamma_j X_i$ for $j = 1,\ldots,5$. The γ parameters give the amount by which the utility parameters for respondents fishing in the DBAY = 1 locations differ from the baseline parameters for respondents fishing in the DBAY = 0 locations. The results for the constrained model with these five extra parameters appear in Table VI. Four of the differences, γ_j, are individually statistically significant at the 10% level according to their asymptotic t-test statistics, but an LR test for the joint contribution of the five additional parameters is only 7.62, which does not allow rejection of the hypothesis that the γ parameters are jointly zero. In this *subsample* of the full dataset, then, there appears to be no statistically significant systematic difference in respondents' utility functions according to the bays in which they fish.

As a second experiment with generalization of the basic quadratic utility model, one can explore the effects of allowing the utility parameters to vary continuously with the level of a sociodemographic variable. In Cameron, Clark, and Stoll (1988), it was found that the Census proportion of people in the respondent's zip code who report themselves as being of

Table VI

Jointly Estimated Models; Utility Parameter Restriction; Correlated Errors; With Heterogeneous Utility Functions[a]

Coefficient	Homogeneous Preferences γ Coeffs = 0	γ Coeff on Bay System Dummy Var.	γ Coeff on Proportion Vietnamese	γ Coeff on Prop. in Farm, Fishing, or Forestry
Shift Variable (X):	-	(DBAY)	(PVIET)	(PFFF)
β_1 (z)	3.309 (8.237)	2.622 (5.320)	2.897 (2.761)	5.082 (5.813)
β_2 (q)	0.1192 (19.55)	0.1140 (13.94)	0.1195 (14.87)	0.1276 (15.92)
β_3 ($z^2/2$)	-0.1167 (-1.836)	0.1138 (2.398)	0.1210 (0.3711)	-0.6539 (-2.790)
β_4 (zq)	0.002579 (2.006)	0.003574 (1.997)	0.003829 (1.800)	-0.0007381 (-0.3813)
β_5 ($q^2/2$)	-0.006837 (-22.80)	-0.006856 (-16.51)	-0.007125 (-21.84)	-0.006743 (-19.79)
γ_1 (zX)	-	2.115 (2.427)	96.64 (0.7534)	-156.0 (-3.533)
γ_2 (qX)	-	0.02276 (1.833)	-0.08279 (-0.09106)	-1.036 (-4.028)
γ_3 ($z^2X/2$)	-	-0.6283 (-2.458)	-58.89 (-1.467)	54.48 (3.027)
γ_4 (zqX)	-	-0.005203 (-1.790)	-0.3573 (-1.395)	0.4691 (4.369)
γ_5 ($q^2X/2$)	-	0.00005803 (0.09560)	0.08352 (6.583)	-0.004925 (-0.3458)
υ	16.01 (81.98)	15.99 (82.01)	15.95 (81.93)	15.99 (81.97)
ρ	0.2315 (9.086)	0.2297 (9.011)	0.2302 (8.971)	0.2312 (9.167)
Max. logL	-15708.17	-15704.36[b]	-15693.97[c]	-15703.99[d]

[a] asymptotic t-test statistics in parentheses.

[b] LR test of parameter restrictions = 7.62 ($\chi^2(5)$ critical value = 11.07); cannot reject "geographical homogeneity" for this partitioning.

[c] LR test = 28.40; readily rejects homogeneity.

[d] LR test = 8.32; cannot reject homogeneity.

Vietnamese origin seemed to be influential in a wide range of models. Table VI demonstrates that the PVIET variable, interacted so as to shift the parameters of the utility function, does indeed make a statistically significant difference to the overall fit of the model and to the parameters of the utility function.

The third experiment exploits the possibility that individuals who live in zip codes wherein a larger proportion of people work in farming, fishing or forestry (PFFF) are likely to have systematically different tastes with respect to recreational fishing. While four of the five additional parameters in this model are individually statistically significant, a joint test of their contribution cannot reject homogeneity.

A visual example of the effect of allowing for heterogeneity with respect to the PVIET variable is displayed in Figures 4a and 4b. As benchmark levels, PVIET=0 and PVIET=.02 are selected. (Maximum PVIET in the sample is 0.0649). The higher the proportion of Vietnamese ancestry in the respondent's zip code, the greater the curvature of the indifference curves, and the larger the implied equivalent variation for a loss of access to the fishery. Correspondingly, in Figure 4b, the demand curve for the PVIET = 0.02 group is situated considerably further out than that for the PVIET = 0 group.

7. IMPLEMENTING THESE PROTOTYPE MODELS IN OTHER APPLICATIONS

For this paper, which illustrates the joint discrete/continuous choice modeling of preference structures using both hypothetical (CV) and real market demand, the data are relatively good but they are still less than ideal. The specific implications of the fitted models must be judged accordingly. If subsequent contingent valuations surveys are to be conducted with the intent to adopt these methods for analysis, it will be important to ensure that the survey collects several pieces of information as accurately as possible.

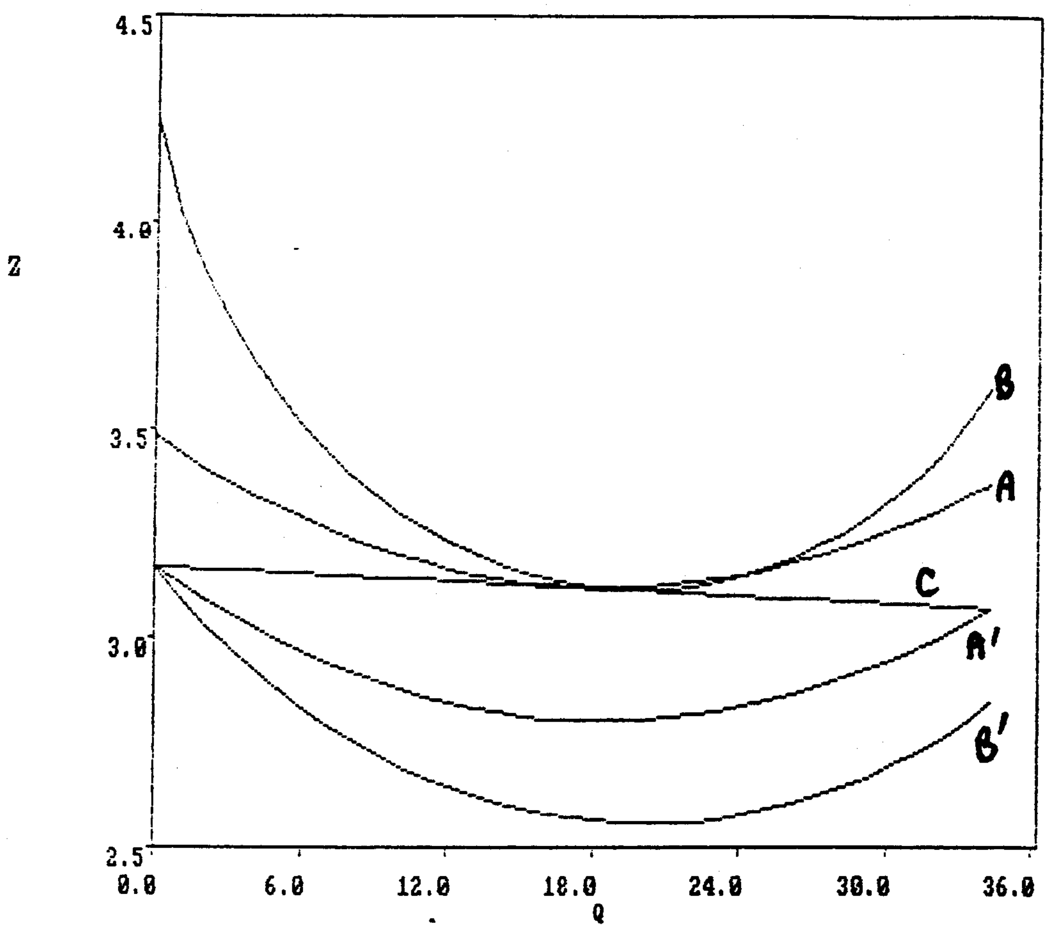

Figure 4a - For respondent with mean income and travel costs, effect of PVIET variable on preference structure: current budget constraint (C), indifference curves if PVIET=0 (A and A'), and indifference curves if PVIET=0.02 (B and B').

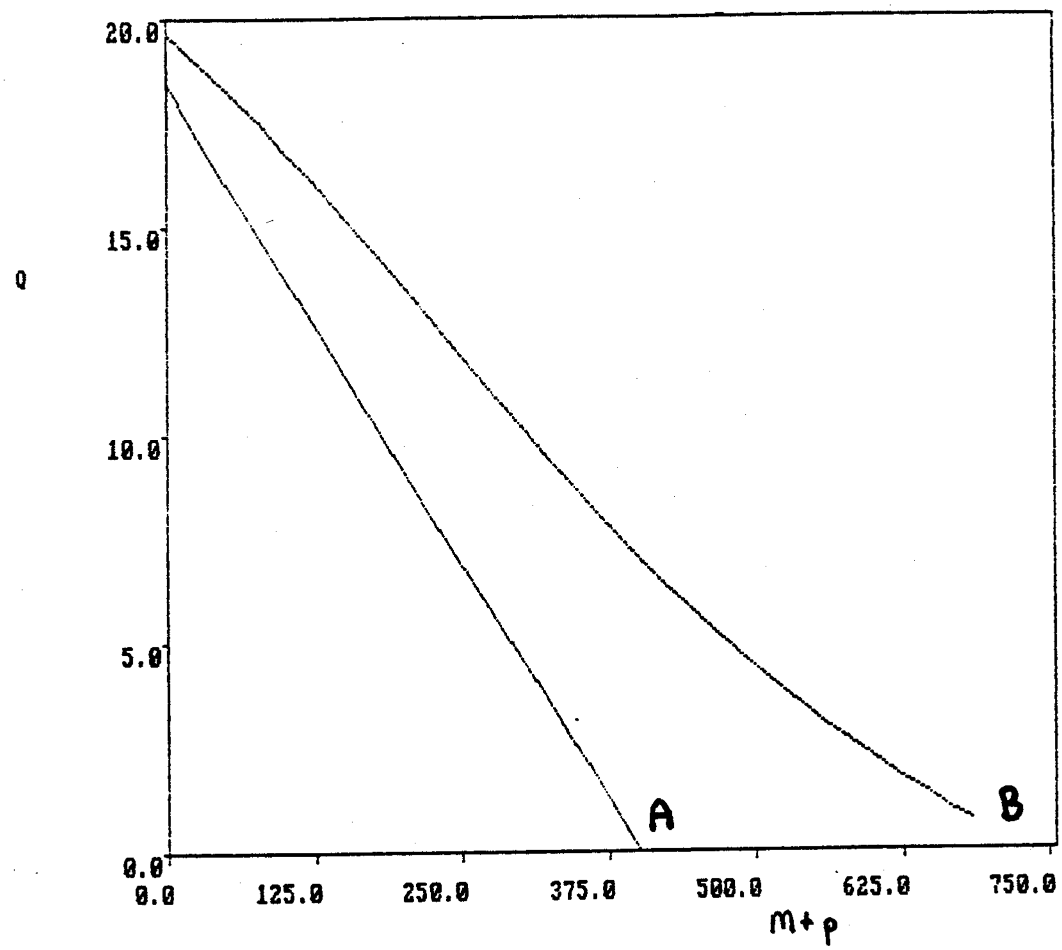

Figure 4b - For respondent with mean income and travel costs, implied demand curves for PVIET=0 (A) and for PVIET=0.02 (B).

First, it is highly desirable to have individual-specific measures of income (and other sociodemographic variables). Census zip code means are helpful, but much information is lost in using group averages as proxies for the true variables. If at all possible, the survey instrument should elicit these data for each respondent.

The contingent valuation question should be phrased so as to make it clear whether the hypothesized change is intended to be a change in relative prices or a lump-sum change in income. This information is vital to the utility-theoretic formulation of the estimating model.

The present survey asks about travel costs for the *current* fishing day. What the model requires is *typical* costs for a *typical* fishing trip. This presumes that individuals fish most of the time at the same location. Much more sophisticated analyses will be required in order to introduce site choice modeling into this framework. (At present, site choice modeling is being pursued in an atheoretic multiple discrete choice framework. Blending the two approaches might have to wait for further computer software and hardware innovations.)

Respondents could be asked specifically about how sure they are concerning their hypothetical responses to the CV question. This information could be incorporated into the weighting scheme for the auto-validation of the CV data.

Option and existence values cannot be captured with the current data set. Selection problems in the assessment of recreation demand have received considerable attention recently (e.g. Smith, 1988). A random sample of households in the target population could be contacted by telephone. If they do not currently consume access days, quantity demanded will simply be zero. Travel costs to relevant sites could still be elicited and appropriate CV questions could be formulated to allow extension of this modeling framework to non-use demands.

8. CONCLUSIONS AND CAVEATS

A *fully utility-theoretic specification* distinguishes this analysis from earlier empirical work on the valuation of non-market resources. By concentrating on identifying the underlying *preference structure* for access days versus all other goods and services, theoretically sound measures of access values (equivalent and compensating variations) can readily be produced.

Several features of the "basic" model should be emphasized. First, it starts from an assumption of quadratic direct utility, which explains the hypothetical contingent valuation responses. Second, the associated non-linear Marshallian demand functions are employed to explain the observed demand decisions by the respondents (a "travel cost" type of model). Third, the corresponding parameters in the utility and the demand functions are constrained to be identical. Fourth, the quantity demanded under the CV scenario is fully endogenized. And finally, unobservable attributes of respondents are allowed to affect both types of responses simultaneously through a non-zero (estimated) error correlation.

The "basic model" forms a minimal prototype for models in a wide range of applications in resource valuation. However, this paper has also described a variety of important extensions--potentially very relevant to subsequent practitioners. Other candidate utility functions have been evaluated and their inadequacies for discrete/continuous auto-validated models have been explained. The basic model can be adjusted to allow for an alternative interpretation of the CV question. "Prior" assumptions about the relative reliability of the hypothetical CV questions can be used to discount the influence of these responses during the process of estimating the utility parameters. It has been shown that if travel cost information is inadequate or unreliable, Engel curves alone can still be used to partially auto-validate the CV responses. Finally, examples have demonstrated that it is straightforward to allow the parameters of the quadratic preference structure to vary systematically with the levels of (exogenous) respondent attributes.

To review the central empirical findings (for these data, in combination with the assumption of quadratic preferences), the "basic model" yields a sample average fitted equivalent variation of $3451 for a complete loss of access to the fishery (with standard deviation of $509). The fitted compensating variation for the same loss is $3560 (standard deviation $535). In contrast, if access days for each individual were restricted by only 10%, the average equivalent variation would be only $35 (standard deviation $5). The implications of the model for "local" variations are probably more reliable, although in this case, the complete loss *is* "within the range of the data" because of the information extracted from the CV responses. This is in sharp contrast to most demand models. Even the auto-validated value estimates, however, are somewhat sensitive to the various extensions of the basic model. For a respondent with sample mean characteristics, Figure 5 summarizes the range of results which can be obtained across the different model specifications. Clearly, the equivalent variation for a loss of access differs systematically across specifications.

One particular caveat should be emphasized. The sample for this application was consciously trimmed along a number of dimensions. Most notably, anyone who reported fishing more than 60 days per year was dropped. When attempting to fit a single utility function to an entire sample, the assumption of identical preferences must be roughly tenable. People who fish

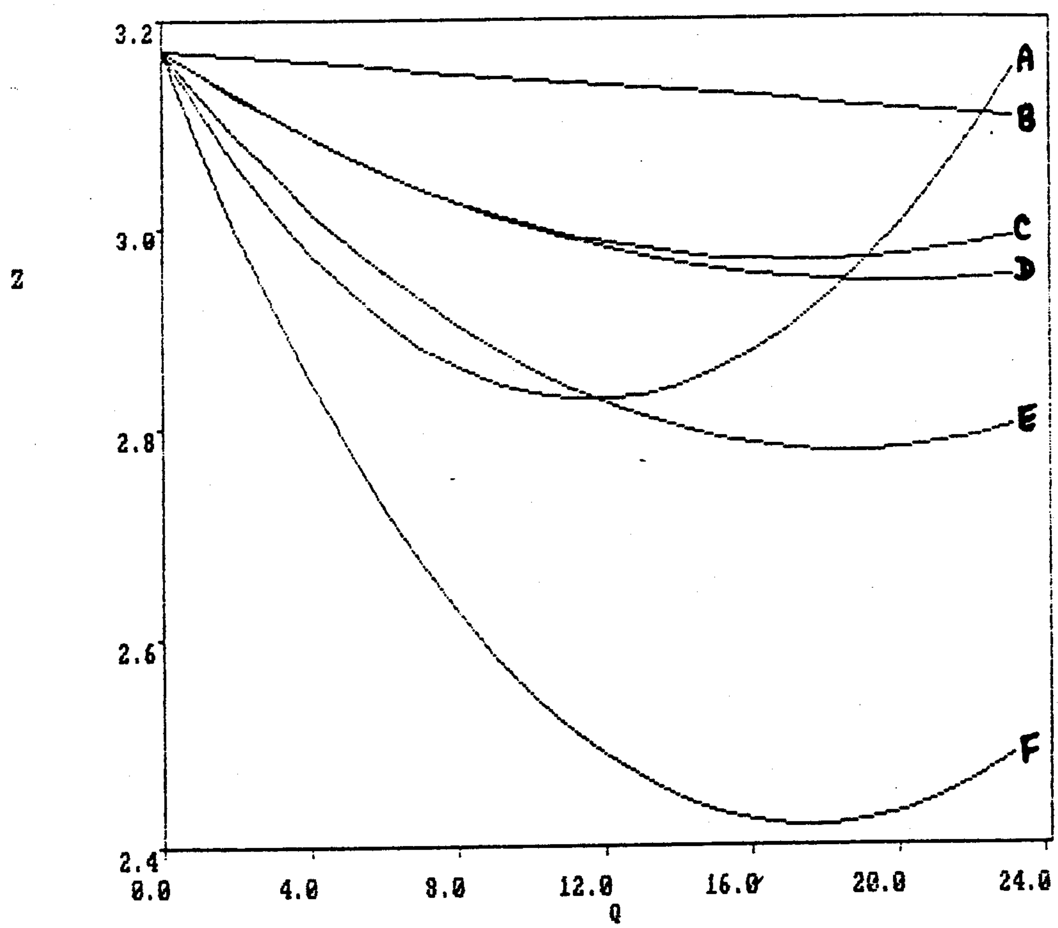

Figure 5 - For respondent with mean income and travel costs, comparison of indifference curves associated with exclusion from the resource for a range of different models. Common current budget constraint (B), indifference curves based upon: CV portion of unrestricted model (A), model ignoring travel costs (C), demand portion of unrestricted model (D), basic model (E), and model using price change interpretation of CV question (F).

200 days per year probably have fundamentally different preferences. With enough detailed information about the exogenous sociodemographic attributes of these individuals which might account for these differences, one could accommodate broad heterogeneity. This survey, however, provides little such information. In order to highlight the capabilities of the model (without obscuring the relationships due to unrecognized heterogeneity), it is necessary to "disenfranchise" some extremely avid anglers. Consequently, if these average values are scaled up to the population of anglers, the total will probably *underestimate* the true value of the fishery. Fortunately, with more detailed surveys (and future generations of computing hardware and software), more comprehensive models will certainly be practicable.

Overall, this research has demonstrated that it is indeed *feasible*, and probably highly desirable, to employ referendum contingent valuation data in the context of a fully utility-theoretic model whenever the quality of the data justify such an effort. (Unfortunately, such data sets are extremely rare at present.) These results also demonstrate that "auto-validation"--forcing contingent valuation utility parameter estimates to be consistent with observed demand behavior--can have a substantial effect on the estimated preference structure, the implied demand functions, and ultimately on the apparent social value of the resource or public good.

It has also been demonstrated that jointly estimating the discrete/continuous choices of respondents *without* auto-validation allows a rigorous statistical check of the consistency of the hypothetical CV responses with demonstrated real market decisions (conditional on the functional form chosen for utility). Previous validation studies have relied on entirely separate models for CV data and other types of data, such as travel cost information or market experiments. Often, completely different samples of respondents have been used. This allows comparisons of *point* estimates of value, but precludes any statistical assessments of the degree of similarily between the results. In contrast, standard likelihood ratio tests are permitted by the auto-validated models presented here. For this sample, the hypothetical data would have to be downweighted drastically before the two sets of utility parameters become statistically indistinguishable. In other cases, however, *consistent* responses under the real and hypothetical scenarios may be readily accepted even with equal weights on the two types of information.

REFERENCES

Agnello, R.J. (1988): "The Economic Value of Fishing Success: An Application of Socio-Economic Survey Data," manuscript, Department of Economics, University of Delaware.

Anderson, L.G. (1980): "An Economic Model or Recreational Fishing," unpublished manuscript, University of Delaware, December.

Bishop, R.C., and T.A. Heberlein (1979): "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics* **61**, 926-930.

Bishop, R.C., T.A. Heberlein, and M. J. Kealy (1983): "Contingent Valuation of Environmental Assets: Comparisons With a Simulated Market," *Natural Resources Journal* 23, 619-633.

Brookshire, D.S., M.A. Thayer, W.D. Schulze, and R.C. D'Arge (1982): "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches," *American Economic Review* 72, 165-177.

Cameron, T.A. (1988): "A New Paradigm for Valuing Non-market Goods Using Referendum Data: Maximum Likelihood Estimation by Censored Logistic Regression," forthcoming, *Journal of Environmental Economics and Management*.

Cameron, T.A., and M.D. James (1987): "Efficient Estimation Methods for 'Closed-Ended' Contingent Valuation Surveys," *Review of Economics and Statistics* **69**, 269-275.

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

CRC (1981): *CRC Standard Mathematical Tables* (26th edition) W.H. Beyer (ed), Boca Raton: CRC Press, Inc.

Donnelly, D.M., Loomis, J.B., Sorg, C.F., Nelson, L.J. (1985): "Net Economic Value of Recreational Steelhead Fishing in Idaho," Resource Bulletin, United States Department of Agriculture, Forest Service, Fort Collins, Colorado 80526.

Dubin, J.A., and D. McFadden (1984): "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," *Econometrica* **52**, 345-362.

Hanemann, W.M. (1984): "Discrete/Continuous Models of Consumer Demand," *Econometrica* 52, 541-561.

Hanemann, W.M. (1984): "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses," *American Journal of Agricultural Economics*, **66**, 322-341.

Huppert, D.D. (1988): "An Examination of Nonresponse Bias and Divergence Among Value Concepts: An Application to Central California Anadromous Fish Runs," Draft report, National Marine Fisheries Service, Southwest Fisheries Center, La Jolla, CA 92037 (May 26, 1988).

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

McConnell, K.E. (1988): "Models for Referendum Data," unpublished mimeograph, Department of Agricultural and Resource Economics, University of Maryland.

McConnell, K.E. (1979): "Values of Marine Recreational Fishing: Measurement and Impact of Measurement," *American Journal of Agricultural Economics* 61, 921-925.

McConnell, K.E. and V.J. Norton, (undated) "An Economic Evaluation of Marine Recreational Fishing: A Review," Department of Resource Economics, University of Rhode Island, Kingston, Rhode Island 02881.

McConnell, K.E. and I.E. Strand (1981): "Some Economic Aspects of Managing Marine Recreational Fishing," in *Economic Analysis of Fishery Management Plans*, L.G. Anderson (ed.), Ann Arbor Science Publishers, Ann Arbor, Mi, 245-66.

Mitchell, R.C., and R.T. Carson, *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Resources for the Future, Washington, D.C. (forthcoming, 1987).

Morey, E.R. and Rowe, R.D. (1985): "The Logit Model and Exact Expected Consumer's Surplus Measures: Valuing Marine Recreational Fishing," unpublished manuscript, Department of Economics, University of Colorado.

Morey, E.R., and W.D. Shaw (1986): "The Impact of Acid Rain: A Characteristics Approach to Estimating the Demand for the Benefits from Recreational Fishing," Department of Economics, University of Colorado, Boulder, CO 80903.

Rowe, R.D., E.R. Morey, A.D. Ross, and W.D. Shaw (1985): "Valuing Marine Recreational Fishing on the Pacific Coast," Report prepared for the National Marine Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, California.

Samples, K.C., and R.C. Bishop (undated): "Estimating Social Values of Sport Caught Fish: A Suggested Approach," unpublished manuscript, Department of Agricultural and Resource Economics, University of Hawaii-Manoa, Honolulu, HI 96822.

Samples, K.C., and R.C. Bishop (1985): "Estimating the Value of Variations in Anglers' Success Rates: An Application of the Multiple-Site Travel Cost Method." *Marine Resource Economics*, 55-74.

Schulze, W.D, R. D'Arge, and D.S. Brookshire (1981): "Valuing Environmental Commodities: Some Recent Experiments," *Land Economics* **57**, 151-172.

Sellar, C., J.R. Stoll and J.P. Chavas (1985): "Validation of Empirical Measures of Welfare Change: a Comparison of Nonmarket Techniques," *Land Economics* **61**, 156-175.

Sellar, C., J.P. Chavas, and J.R. Stoll (1986): "Specification of the Logit Model: the Case of Valuation of Nonmarket goods," *Journal of Environmental Economics and Management* **13**, 382-390.

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

Vaughn, W.J., and C.S. Russell (1982): "Valuing a Fishing Day: An Application of a Systematic Varying Parameter Model," *Land Economics* **58**, 450-463.

APPENDIX I

An Alternative Model: Stone-Geary Utility Function

Empirical research on consumer decisions has sometimes employed the Stone-Geary utility function and its corresponding "linear expenditure system" demand equations. This specification may at first seem attractive. However, it seems to be appropriate only when one is considering decidedly *interior* consumer optima. This appendix explains why the Stone-Geary function will be inappropriate for the CV scenario posed in the survey used for this study. In this case, the utility function would be:

$$U\ (z,q) = (z - \beta_1)^{\beta_2}\ (q - \beta_3)^{\beta_4}$$

The corresponding demand for fishing days will be given by:

$$q = \beta_3 + (\beta_4/p)\ [Y - \beta_1 - \beta_3\ p]$$

where the price of the composite good, z, has again been normalized to unity.

This utility function is not linear in parameters, so initial estimates cannot be obtained via a conventional maximum likelihood probit package. But there is a bigger problem, stemming from the necessity of considering utility levels for zero days of access. In particular, the systematic portion of the utility difference function, which would form the non-linear "index" function for the discrete choice portion of the model, would take the following form:

$$\Delta U = [(Y\text{-}M\text{-}T) - \beta_1]^{\beta_2}\ [q - \beta_3]^{\beta_4} - [\ Y - \beta_1\ ]^{\beta_2}\ [\ - \beta_3\ ]^{\beta_4}$$

The problem for estimation stems from the last term. The coefficient β_4 is often fractional. Attempting to take the β_4-root of a negative number can be expected to create difficulties. FORTRAN, in particular, computes the values of terms such as these by first taking logs. The logarithm of a negative number is undefined. Furthermore, the usual interpretation of β_3 is that is represents "subsistence" consumption levels of commodity q, so negative values of the parameter itself are unlikely to result, or to be defensible intuitively, if they do. As expected, in attempts to estimate this model using the data employed in the rest of this study, the algorithm persistently failed.

NOTES

[1] These statistics are cited in an article in *Forbes*, May 16, 1988, pp. 114-120.

[2] These data do not allow accurate imputation of the opportunity costs of travel time.

[3] With better information regarding the prices of these market goods and the actual quantities consumed of each, one could of course "unbundle" access days from these market goods and enter market goods (or an index of these goods) and access days separately in the utility function. This would be highly desirable, because it seems clear that, realistically, market goods are not consumed in fixed proportion with access days. Access days vary in their attributes (e.g. different sites), so market goods consumed will also differ (e.g. different travel costs).

[4] I.e. if you do not "gain admission to the market" by paying tax T, then no amount of income forgone can legally be traded (down the usual budget constraint) for positive access days. The budget constraint is effectively vertical at zero days until access is purchased.

[5] One could, of course, shift the utility surface one unit towards the origin along the dimension of each good by adding one to each quantity within the functional form for the translog direct utility. However, when the direct utility function, rather than the indirect utility function, takes on a translog functional form, the associated Marshallian demand functions are awkward to derive; they are even more awkward if the function is additively shifted.

[6] Huppert's payment card contingent valuation responses are treated as a continuous variable, so that the joint estimation of the utility and demand parameters can be accomplished via simultaneous non-linear least squares algorithms.

[7] In-person CV surveys typically sample only current users of the resource. When access price increases (or simply positive access prices) are being contemplated, this does not pose much of a problem. However, when projected scenarios involved improved resource attributes, one must really survey potential users as well as current users to elicit an accurate measure of aggregate demand responsiveness.

[8] An intercept term could be included in ΔU to allow for the possibility of a discontinuity in the utility surface between $q = 0$ and $q > 0$. If utility levels are systematically higher when any non-zero level of resource access is consumed, a coefficient β_0 would reflect this. Such a vertical displacement in the utility surface does not affect any of the first order conditions for constrained optimization, but does complicate somewhat the interpretation of the social cost of prohibiting access to the fishery. A full set of results was initially obtained for models incorporating such an intercept.

[9] Subsequent research should concentrate on more-flexible representations for utility, to minimize the risk that the functional form accounts for the parameter inconsistency, rather that the hypothetical nature of CV responses.

[10] Unfortunately, the duration of each trip is unknown, so it must be assumed that the majority are one-day trips, which may or may not be entirely plausible. Here, the term "trip" is used synonymously with "fishing day."

[11] In these demand models, it will be easy to test whether the income effect is in fact zero.

[12] Providing that the assumption of $q^0 = q^1$ is tenable.

[13] Huppert (1988) is also careful to substitute the endogenous formula for quantity into his utility specification.

[14] The standard deviation υ can be estimated explicitly, whereas σ is "bundled" with the β parameters in the linear index of the discrete choice portion of the model. Fortunately, since the β's in the demand model are ratios of the βs in the discrete choice model, the unidentified σ parameter causes no mischief.

[15] Huppert (1988) does not allow for correlated errors in his joint estimation of log-linear travel cost demand parameters and the corresponding utility parameters.

[16] If the estimated value of the error correlation, ρ, is substantial and statistically significant, one probably ought to generalize the specification, if possible, to accommodate systematic heterogeneity across respondents. Section 6 will address this issue.

[17] For the single individual with average characteristics in Figure 1, this quantity would be determined by taking the parallel downward shift in the budget constraint which would leave the new constraint just tangent to the lower indifference curve.

[18] The computed equivalent variation, as a function of α, is convex when viewed from below.

[19] Intuitively, one would expect that additional information would be required to "identify" the parameters in this case, and optimization failure corroborates this.

[20] This search could readily be "automated" by embedding the two optimization problems (unconstrained and constrained) within a loop which checks the likelihood ratio test statistic for each weighting scheme, compares it to the critical value of the $\chi^2(4)$ distribution and chooses a new weighting value until the value of the test statistic just equals the critical value for some pre-selected level of significance.

[21] This will *not* be a general result for these models.

[22] Earlier models employed an intercept term in the utility-difference function. For those models, the difference between the CV parameters and the travel cost parameters in the unrestricted model became statistically insignificant at the 5% level as the reliability of the CV data was reduced from 1/10 to 1/11.