NON-MARKET RESOURCE VALUATION: ASSESSMENT OF VALUE ELICITATION BY "PAYMENT CARD" VERSUS "REFERENDUM" METHODS

by

Trudy Ann Cameron
University of California, Los Angeles

and

Daniel D. Huppert
Nati‹nal Marine Fisheries Service
Southw‹st Fisheries Center, La Jolla

Discussion Paper #448
Department of Economics
Univeristy of California, Los Angeles
405 Hilgard Ave
Los Angeles, CA 90024-1477

July 1987

July 17, 1987

Please do not cite or quote without authors' permission

NON-MARKET RESOURCE VALUATION: ASSESSMENT OF VALUE ELICITATION BY "PAYMENT CARD" VERSUS "REFERENDUM" METHODS.

by

Trudy Ann Cameron*
Department of Economics, UCLA
405 Hilgard Avenue
Los Angeles, CA 90024-1477

and

Daniel D. Huppert
National Marine Fisheries Service
Southwest Fisheries Center
8604 La Jolla Shores Drive
La Jolla, CA 92308

ABSTRACT

Contingent valuation methods have been shown to be extremely useful for eliciting information about demands for non-market goods. This paper examines the implications of a survey which collects valuation information with a "payment card" vehicle and compares these to the range of results which *would* have been generated if (i.) coarser intervals had been specified on the payment card, or (ii.) a "referendum" format had been used instead. The *true* payment card data are used in both a.) a naive ordinary least squares procedure employing interval midpoints as proxies for the true dependent variable, and b.) a maximum likelihood (ML) procedure which explicitly accommodates the intervals. The ML procedure is also used to compare different degrees of interval coarseness. The *artificial* referendum data are simulated by Monte Carlo experiments. Our empirical example is the valuation of a recreational fisheries enhancement program. We examine the different implications to be drawn from these data, depending upon the estimation method used and upon the quality of the valuation data. In addition to the purely econometric issues which are the focus of this paper, we are also able to offer some insights on the social value of this enhancement program.

* We thank Michael Hanemann, John Loomis, and Jane Murdoch for helpful comments and suggestions.

NON-MARKET RESOURCE VALUATION: ASSESSMENT OF VALUE ELICITATION BY "PAYMENT CARD" VERSUS "REFERENDUM" METHODS.

by

T.A. Cameron and D.D. Huppert

1. Introduction

Contingent valuation methods have now been recognized as extremely useful devices for eliciting demands for non-market goods (see Cummings, Brookshire, and Schulze, 1986, and Mitchell and Carson, 1987, references cited therein). These methods use hypothetical market scenarios posed to individual respondents to discern their likely behavior under different conditions. While we would prefer to elicit a respondent's exact willingness-to-pay, WTP (or willingness-to-accept, WTA), for an increase (decrease) in consumption of a public good, it has been argued that respondents find it very difficult to name a specific sum. To avoid non-response, researchers have found it more fruitful to pose the valuation questions differently.

For telephone or in-person surveys, one alternative is called "sequential bidding." If a respondent indicates a willingness to pay the initial amount offered, a larger amount is proposed. The bidding continues until the interviewer reaches an amount that the respondent will not pay. However, it has been argued that the starting point (i.e. whether it lies above or below actual valuation), can influence the outcome. "Referendum," or "closed ended contingent valuation" (CECV) surveys are therefore often preferred for telephone or in-person interviews. In this format, the respondent is offered just one randomly assigned amount and their yes/no response is recorded. The valuation information in these responses is much more diffuse, so a large number of responses is required to identify the approximate valuation. However, much less stress is placed upon the respondent since the pricing scenario mimics the take-it-or-leave-it market decisions made daily by most of us. Consequently, a higher completed response rate can probably be expected.

In mail surveys, however, an alternative format is feasible. Rather than sequential bidding, the survey questionnaire can be designed to include an ordered set of threshold values, called a "payment card." The respondent is then asked simply to peruse the range of values and to circle the highest (lowest) amount they would be willing to pay (accept). As with oral sequential bidding, it can then be inferred that the respondent's true "point" valuation lies somewhere in the interval between the circled value and the next highest (lowest) option. As with CECV methods, payment cards conserve respondent effort relative to oral sequential bidding. Even a fairly detailed set of thresholds can be visually scanned quite quickly.

Of course, the ideal estimation procedure would involve regression of "point" valuations on a range of economic or sociodemographic variables collected along with the valuation information. Neoclassical microeconomic theory should suggest appropriate candidate functional forms for this underlying "inverse" demand function. Without these point valuations, a simplistic alternative for utilizing the information that the true value lies somewhere in a known interval would be to assign the midpoint of the relevant interval as a proxy for the mean of the variable over that interval and to employ ordinary least squares (OLS) regression using these midpoints as the dependent variable.

There is, however, a considerable statistical literature on the efficient maximum likelihood (ML) estimation of regression models where the dependent variable is only measured on intervals of a continuous scale (see Hasselblad, Stead, and Galke (1980), Burridge (1982), Stewart (1983), and Cameron (1987a)). These more-sophisticated methods are appropriate in this context. One task of this paper is to assess the distortion introduced into contingent valuation estimates due to the inappropriate application of OLS regression methods to payment card interval midpoints.

Other objectives include comparison of the efficiency of payment card estimates of valuation functions with the results which *would have* been obtained (a.) if coarser intervals had been offered, or (b.) if the respondents had been posed closed-ended contingent valuation questions instead. In the latter case, we will assume that *any one* of the payment card values might have been assigned as the single CECV threshold, so Monte Carlo methods are necessary to identify the range of parameter estimates and fitted valuations which might have been attained under CECV methods.

Section 2 of this paper outlines the procedure for maximum likelihood estimation of the "payment card" regression model Section 3 describes maximum likelihood estimation of "referendum" (CECV) models. Section 4 briefly summarizes the dataset. Section 5 addresses the specification of the valuation (WTP) function. Section 6 compares "naive" results from OLS estimation using interval midpoints with those from the more-appropriate ML interval estimation method. Section 7 explores the potential consequences had the questionnaires been designed with coarser payment card intervals. The design of the Monte Carlo exercises is explained in Section 8 and the actual interval estimates are compared to the range of CECV estimates which might have been obtained had a single-value referendum question format been used instead. Section 9 concludes, offering both caveats for the interpretation of the current work and directions for further research.

2. Maximum Likelihood Estimation with "Payment Card" Data

Since previous empirical studies (Cameron and James, 1986, 1987) have indicated that the distribution of valuations is frequently skewed, we will propose that a *lognormal* conditional distribution for valuations should serve as a useful first approximation. If the respondent's true valuation, Y_i, is known to lie within the interval (t_{li}, t_{ui}), then $\log Y_i$ will lie between $\log t_{li}$ and $\log t_{ui}$. It is generally assumed that $E(\log Y_i | x_i)$ is some function $g(x_i, \beta)$, for which a linear form is convenient. If $\log Y_i = x_i'\beta + u_i$ and u_i is distributed normally with mean 0 and standard deviation σ, then we can standardize the range of values $\log Y_i$ occupies and state that:

$$(1) \qquad \Pr(Y_i \in (t_{li}, t_{ui})) = \Pr((\log t_{li} - x_i'\beta)/\sigma < z_i < (\log t_{li} + x_i'\beta)/\sigma)$$

where z_i is the standard normal random variable. Let z_{li} and z_{ui} signify the lower and upper limits, respectively. The log-likelihood function for a sample of n independent observations is then:

$$(2) \qquad \log L = \sum \log [\Phi(z_{ui}) - \Phi(z_{li})]$$

where Φ is the cumulative standard normal density function. Appendix 1 details the formulas for the gradients and the Hessian matrix associated with this likelihood function.[1] In this study, we use the computer package GQOPT (Goldfeld and Quandt) to maximize the necessary likelihood functions with respect to the unknown parameters, β and σ. However, for this interval data model, researchers who have access to the LIMDEP computer package (Greene,

---

[1] These formulas can substantially reduce computational costs for the optimization of the log-likelihood function.

1986) will probably find that the "GROUPED DATA" procedure described there can be used to determine these parameters.

Once the optimal values of β and σ have been attained, it is a simple matter to reconstruct fitted values of the transformed variable log Y. The conditional mean of log Y for any given vector of x variables will be $x_i'\beta$. However, if we wish to retransform back to the estimated conditional distribution for the variable Y, $\exp(x_i'\beta)$ will be the *median* of the distribution of Y rather than the mean.[2] These quantities are a valid measure of the central tendency of the unobserved dependent variable Y. However, we may also wish to compute the *mean* of the Y variable, which is obtained by scaling this quantity by an estimated constant equal to $\exp(\sigma^2/2)$. Either the fitted medians or the fitted means can then be used to compute the weighted average of fitted valuations across the sample.

3. Maximum Likelihood Estimation with Referendum Data

Assume again that the respondent's true valuation is Y_i, and that $\log Y_i = x_i'\beta + u_i$, where u_i is distributed normally with mean 0 and variance σ. Suppose we are considering a WTP scenario: the respondent is offered a single threshold value t_i; if he is willing to pay this amount, we record $I_i = 1$ (if not, $I_i = 0$). Then we can presume that:

$$
\begin{aligned}
(3) \qquad \Pr(I_i = 1) &= \Pr(\log Y_i > \log t_i) = \Pr(u_i > \log t_i - x_i'\beta) \\
&= \Pr(u_i/\sigma > (\log t_i - x_i'\beta)/\sigma) \\
&= 1 - \Phi((\log t_i - x_i'\beta)/\sigma).
\end{aligned}
$$

The log-likelihood function is then:

---

[2] We are grateful to Michael Hanemann for reminding us of this point. For a description of the lognormal distribution, see Hastings and Peacock (1974, p. 84).

$$(4) \quad \log L = \sum \{ I_i \log [1 - \Phi ((\log t_i)/\sigma - x_i'\beta/\sigma)] + (1 - I_i) \log [\Phi ((\log t_i)/\sigma - x_i'\beta/\sigma)]\}.$$

The presence of $\log t_i$ allows σ to be identified so that the underlying valuation function, $x_i'\beta$, can be recovered. (Note that if $\log t_i = 0$ for all i, we have the conventional maximum likelihood probit model.) A full description of the elements of the gradient and the Hessian has been relegated to Appendix 2. We know of no packaged computer program which can estimate this model directly, but it has been shown elsewhere (Cameron and James, 1987) that conventional probit algorithm parameter estimates for appropriate specifications can be transformed to yield point estimates for β and σ, and standard errors for these coefficients can be computed readily by Taylor's series approximation.

Again, if we desire the fitted *median* valuation Y for a given x vector, we will use the fitted values of β to compute $\exp(x_i'\beta)$; if the *mean* of Y is preferred as a measure of central tendency, we must use $\exp(x_i'\beta)\exp(\sigma^2/2)$.

4. The Data

Our data are drawn from the NOAA National Marine Fisheries Service "Saltwater Recreational Fishing Survey," of California's San Francisco Bay Area. Detailed summaries of the responses, as well as a copy of the questionnaire, have been relegated to extensive supplementary documentation (Thompson and Huppert, 1987). Only a brief summary of the pertinent responses will be provided here.

The crucial valuation question examined in this paper was worded as follows: "What is the MOST you would be willing to pay each year to support hatcheries and habitat restoration that would result in a doubling of current salmon and striped bass catch rates in the San Francisco Bay and ocean area if without these efforts your expected catch in this area would remain at current levels? (*Circle the amount*)." The listed values were $0, $5, $10, $15, $20, $25, $50, $75, $100, $150, $200, $250, $300, $350, $400, $450, $500, $550, $600, and "$750 or more." Respondents who circled $0 were also asked: "Did you circle $0 because you feel this change has no value to you?" If a respondent replies "no" to this question, we assume that this answer indicates that they value the change by more than $0 but by less than $5.

Descriptive statistics for the usable portion of the sample are given in Table I. Note that we do not attempt to compensate for the measurement error inherent in the income (INC) variable because of its categories; this issue is beyond the scope of this study. The pair of dummy variables (STRIP and BTRIP) distinguishes between anglers who sought to catch only one particular species and anglers who sought either species, without preference. Ability was rated on a subjective 5-point scale, with 1 being "novice" and 5 being "expert." These categories were aggregated into "below intermediate," "intermediate," and "better than intermediate." The two lowest categories was taken as the base case and the pair of dummy variables (ABIL2 and ABIL3) capture the higher skill levels. If the respondent indicated that they owned or operated a boat that could be used for saltwater fishing, the dummy variable OWNBOAT takes on a value of 1.

We will maintain the hypothesis that current catch rates will influence respondent's basic valuation of the fishery, and should therefore be expected to affect their valuation of a doubling of the catch rates. The survey was conducted by mail, and asked for retrospective information on each respondent's actual catch for their third-last, second-last, and most recent fishing trips (if at least one trip was taken over the last twelve months). We are focusing in this study on fishing trips for which the target species was exclusively salmon, exclusively striped bass, or either salmon or striped bass. The catch data are in the form of "catch per trip," and are summarized by the variables C1SAL, C2SB, C3SAL, and C3SB, described in Table I. Observed catch may not be fully exogenous (due to the fact that respondents' fishing abilities and choice of site may affect both their average catch and their valuation), but we control to a considerable extent for differing abilities with the subjective ABIL2 and ABIL3 dummy variables. Neither do observed average catch rates for the last three fishing trips by a particular individuals necessarily provide a good estimate of their *expected* catch. We would anticipate that this expected catch has a greater impact upon the demand for fishing days than do recent actual catch rates.

5. Relation of the Valuation (WTP) Function to the Inverse Demand Function

In order to use the econometric models described in sections 2 and 3, we must first decide upon plausible theoretically consistent specifications for the presumed underlying valuation function. This function has the general form $Y_i = g(x_i, \beta)$, where $g(x_i, \beta) = x_i'\beta$ in the simple case. It is the task of the investigator to decide upon which of the available variables (or which transformations of these variables) belong in the vector x_i.

To remain loyal to the neoclassical economic theory of constrained utility maximization, we would prefer to employ a specification for $g(x_i, \beta)$ which is consistent with a form for the valuation function which can be interpreted as an inverse Hicksian-compensated demand function (derived from some utility function which has desirable properties). In Cameron (1987b), however, it has been argued that earlier analyses of contingent valuation data (especially referendum data) have placed too much emphasis upon the underlying utility function (see in particular Hanemann, 1984). Since it is *not* necessary at any time actually to estimate the utility function, we need not be limited to the awkward functional forms for the inverse demand function which correspond to convenient (tractable) linear-in-parameters specifications for the "utility difference" function.

In many cases, almost any specification for an inverse demand function for which sufficient variables are available (either *ad hoc* or utility-theoretic) can be estimated directly, regardless of whether the data are collected by payment card or by referendum survey. Unfortunately, our valuation question does not elicit information on the *height* of the inverse demand curve itself.[3] To be consistent with the way the question is worded in the survey, we must consider the estimated value, either $\exp(x_i'\beta)$ or $\exp(x_i'\beta)\exp(\sigma^2/2)$, to be a measure of the increase in total surplus due to whatever *vertical shift* in the demand function is to be expected from a doubling of current overall salmon and striped bass catch rates.[4]

It is important to consider carefully the implied bivariate relationship between anglers' willingness to pay (WTP) for fishing days (before and after a doubling of the catch rate) versus number of days spent fishing (i.e. holding all other explanatory variables constant). If we were to assume that this change in value depended linearly upon the number of fishing days, we would be imposing a *parallel* upward shift of the "demand" curve as a result of doubling the fish stock . Only then would the change in the area between the curves be proportional to a change in number of days at which it is calculated.

---

[3] We could conceivably attempt to utilize information on travel costs to identify the position of the current demand curve, but this procedure is fraught with difficulties arising from the ambiguity in assigning actual time costs faced by anglers. This task is held over for subsequent analyses.

[4] Note that this is expressly not a doubling of the number of fishing days, the "quantity" of the good for which demand is being modeled, but a doubling of one of the factors which is argued to "shift" this demand curve.

Instead, we employ a log-linear transformation of this area as our implicit dependent variable. This allows the shift to be non-parallel.

It is important to keep in mind that the fisheries enhancement program being proposed must really be interpreted as an increase in the stock of fish available for catching. The question is posed in terms of catch rates *overall*, not in terms of the *individual's* catch rate. There is only a loose correlation between fish stocks and individual catch rates (and in many instances, the individual's catch is currently zero). Hence, in this study, we do not pursue translation of "WTP for a doubling of overall catch rates" into "WTP for an additional fish"--a measure which has been sought in other studies.[5] The ultimate objective in a project like this one is a measure of the social value of the enhancement project (i.e. the social value of having twice as many fish out there). This figure could presumably be compared to the project's overall cost to assess its advisability.

Unfortunately, as is frequently the case in empirical applications, we do not have enough observations or sufficiently accurate proxy variables to support a fully utility-theoretic specification which also controls for observable sociodemographic heterogeneity among the households. Preliminary OLS regression models using the midpoints of each of the valuation intervals were used to determine the most likely set of explanatory variables as well as to assess whether linear or log-linear specifications might be appropriate. Among simple forms, it seems that a log-linear specification is most

[5] In auxiliary experiments, we have segregated respondents according to their target species. For each of these groups, we have transformed the midpoint of the valuation interval by dividing through by the respondent's average catch, if catch was non-zero. This makes the very strong assumption that if overall catch rates double, so will this respondent's catch. This procedure generates a new dependent variable: value of increasing expected catch by one fish. In these models, however, this value still depends significantly on the number of fish caught.

appropriate and that the variables TRPS, log(INC), STRIP and BTRIP, ABIL2 and ABIL3, OWNBOAT, and the set of catch-per-trip variables are the most appropriate (first-round) explanatory variables in *ad hoc* specifications. Note that, initially, we retain the full set of each group of related variables in a "full" model, despite the fact that some members of these sets may have statistically insignificant coefficients. When we find that alternative estimation techniques suggest that different subsets of these variables make statistically significant contributions to explaining valuation, we provide the full model and a "brief" model in each case.

6. Comparison of OLS Midpoint Estimates and Maximum Likelihood Estimation using Interval Data

Using the midpoints of each payment card valuation interval in conventional OLS estimation requires the assumption that the midpoint of each interval is an adequate proxy for the conditional mean of the dependent variable over that interval. Depending upon the coarseness of the grouping interval, this assumption may or may not be tenable. Our first task, therefore, is to compare the point estimates from the OLS midpoint method with those from the maximum likelihood (ML) estimation procedure (described in section 2) which explicitly accounts for the interval nature of the dependent variable.

Table II gives parameter point estimates for the two different estimation methods (for each of two specifications), along with the OLS t-ratios and the ML asymptotic t-test statistics.[6] As anticipated (based on earlier work by Cameron, 1987a), using the interval midpoints as proxies for

[6] Normalized weights have been incorporated; these scale in the influence of each observation to reflect population frequencies of the different types of respondents according to a cross-tabulation of each group by county of origin and by frequency of fishing trips over the last 12 months.

the true values of the dependent variable in a log-linear normal specification results in moderate distortion of the point estimates.

Since the results for all of the models to be estimated here are qualitatively similar, it is worthwhile to look briefly at the plausibility of our parameter estimates. Examination of the OLS midpoint estimates shows that the number of trips has a very small but significant effect on the log of WTP for a doubling of catch rates. This means that the "before" and "after" demand curves *diverge* slightly as the number of days increases. The doubling of the catch would also appear to be a "normal good" for which demand is inelastic with respect to income.

Concerning the target species, the base case is trips upon which *either* species was sought. If the angler sought only salmon on his last three trips, WTP is lower by about 46% (in the "brief" model). If only striped bass were sought, WTP is lower by about 58%. On the surface, it would appear that anglers who care only about one of these two species are only about *half* as willing to pay for enhancement measures which would double the catch of *both* species. Only the highest ability rating enters significantly. Anglers with above-average ability have WTP about 56% higher. For serious anglers, it is possible that the catch figures more prominently among the "bundle" of utility-generating activities that make up a fishing day. Boat ownership enters significantly, contributing to a decrease of over 40% in WTP. This would seem to imply that the easier it is for an angler to increase his number of fishing trips, the less willing is he to pay for higher catch rates per trip. This is plausible.

Among the average catch rate variables, only the salmon catch rate (on exclusively *salmon*-fishing trips) significantly affects WTP. If, for a given respondent, the average salmon catch rate on past salmon-fishing trips was higher by one fish, it would not be surprising that this individual would be more satisfied with current catch rates and therefore less willing to pay for a doubling of overall catch rates (by approximately 17%).

In contrast, we can compare the full models estimated by OLS and ML. Of these two methods, the maximum likelihood method is unequivocally the superior analytical technique, since (aside from the usual distributional assumptions) it imputes no more information about the dependent variable than is actually contained in the interval responses. In this example, while we do not know the true underlying parameter values, the evidence from the full models suggests that the OLS midpoint method overstates the income elasticity of demand for the enhancement project. It also overstates the decrease in WTP due to focus on a particular target species, overstates the effects of ability, boat ownership, and the average catch on salmon-fishing trips. Furthermore, the STRIP and OWNBOAT variables which appear to be statistically significant determinants of WTP in the OLS model lose their apparent significance when the more-valid ML interval technique is employed.[7] The maximized value of the log-likelihood function is provided for each estimation method, but bear in mind that these values are not comparable since the models are non-nested and have very different dependent variables.

As noted in sections 2 and 3, we can use either the median or the mean of the fitted conditional distribution of valuations to represent the most likely value of the unobserved dependent variable for each respondent. The skewness of the lognormal probability density function will cause the median to be smaller than the mean. While Table II describes the different point

---

[7] While the differences for some coefficients appear negligible, bear in mind that these are *percent* measures (in this log-linear specification); also, when scaled up to population values, these small differences can become very large in the aggregate.

estimates generated by the two estimation methods, we might also be interested in the implied derivatives of the valuation difference, Y, with respect to each explanatory variable. Since the implicit dependent variable is in log form, each estimated coefficient must be multiplied by the fitted value of Y_i to yield a derivative for each observation. Use of the median for this quantity gives the results reported in Table III(a); the mean yields Table III(b). We compute the value of each derivative for each respondent and then compute the mean and standard deviation of these derivatives across the sample.[8]

A few comments on the discrepancies between OLS and ML derivatives are warranted. In the "full" models, OLS with midpoints appears to overstate substantially the decrease in value when the respondent's recent fishing trips have targeted salmon exclusively. The effect of above-average ability is also overstated by a large margin (about $7.00, on average, using medians, and by much more when means are used), as is that of boat ownership (by about $8.00 using medians). These differences reflect both differences in the individual point estimates and the differences in the fitted valuations which result when a different β vector leads to a different fitted value for $x_i'\beta$ (used in computing both the median or the mean).

In addition to assessing different methods for achieving the parameter point estimates and the implied derivatives, it is useful to compute the implications of the fitted model for the empirical question at hand, namely the valuation of changes in the stock size. The lower portions of Tables III(a) and III(b) provide within-sample weighted averages for the marginal distribution of median and mean valuations, respectively. Relative to the

---

[8] These means are also weighted to reflect population frequencies of respondent types.

more-appropriate ML interval estimation method, therefore, it also seems that the OLS midpoint technique substantially underestimates the marginal mean of the distribution of fitted WTP values--by almost 20% in the case of either medians or means. This large a difference *could* have a substantial influence on policy decisions.

7. Comparison of Actual Interval Results with those for Coarser Intervals

The finer the intervals listed on the payment card, the more difficult it becomes for a respondent to decide exactly which interval contains their actual willingness to pay. In this section, we consider the effects upon the parameter point estimates and marginal mean valuations of imposing coarser and coarser groupings. Throughout, we presume that if a respondent indicates a particular interval when there are twenty to choose from, they would have indicated the corresponding larger interval when these finer intervals are aggregated. Two counterfactual interval specifications are explored.

a. Coarser Intervals

First, the existing intervals are collected in pairs, then triples, then quadruples. This allows us to explore what would have been the consequences--for the point estimates, the implied derivatives, and the mean fitted WTP--had the payment card been designed with coarser and coarser intervals. Parameter estimates for this experiment are contained in Table IV(a). Broader intervals make it easier for the respondent to decide in which interval his true valuation lies, but clearly, there is a tradeoff in terms of the loss of information for estimation purposes. Cameron (1987a) demonstrates that as interval widths increase in a log-linear normal specification, it is unfortunately not possible to systematically sign the bias introduced into the parameter estimates--these biases oscillate. This tendency is observed as one scans across the columns of Table IV(a). Implied derivatives and means of the fitted values of WTP using medians and mean appear in Tables IV(b) and IV(c) respectively. Note that the implied derivatives seem to be quite sensitive to the coarseness of the intervals. This point should be kept in mind in designing payment card intervals. Even more disturbing is the apparent tendency for the overall marginal mean of the fitted valuations to increase systematically with interval width. For an unscrupulous researcher, it seems that the apparent social value of the resource in question could in some cases be increased by offering relatively coarser payment card intervals.

b. Fewer Upper Intervals

In the second counterfactual experiment, the existing intervals are retained for the lower levels, but the open-ended upper interval changes from >750 to >300, then to >200, then to >100. Table V(a) gives the point estimates and asymptotic t-statistics. The implied derivatives and the marginal means of fitted WTP corresponding to these simulations, using medians and means, appear in Tables V(b) and V(c), respectively. The footnotes to Table V(a) indicate how many observations are affected by the aggregation of the upper intervals. These experiments show the considerable influence upon the results of the study wielded by "extreme" responses. In particular, a decision not to discriminate so finely among the upper intervals results in considerable "damping" of the implied derivatives and of the overall marginal mean of fitted WTP. Again, it seems that an unscrupulous researcher could sometimes have a substantial effect upon the estimated total value of the resource by tailoring the upper intervals appropriately.

8. Comparison of ML Interval Estimates with Simulated Referendum Data

Valuation data in this survey were in actuality collected by means of payment card, but it would be interesting to assess how well a referendum format would have performed. With the referendum format, less "information" is available, so we would expect that more observations would be necessary to achieve the same level of accuracy in the estimation of the coefficients. However, less effort (introspection) is required of the respondent, which has the potential for increasing the completed response rate. In this section, we take the payment card responses and use them to construct the "yes/no" responses we would have expected if the valuation question had been posed in the form of a single threshold value.[9]

Given the format of the payment card actually used, there are twenty different "thresholds" we might have assigned for each respondent. Monte Carlo techniques are therefore required to generate information concerning the range of possible estimates that might have been achieved under the referendum format. Since more thresholds are typically offered at the lower end of the value spectrum (to increase the resolution in the range where it is suspected that a large proportion of the valuations lie), we opt to generate threshold values from frequency distributions similar to the observed frequencies for the valuations implied by our respondents.

Two alternative underlying distributions were assumed for the single thresholds to be "offered" to each respondent in our simulations of the referendum questions. The first distribution mimics the distribution of

[9] Each referendum threshold divides the range of possible valuations into two "intervals." However, the referendum approach differs from the payment card method in that these thresholds are varied across respondents. With the payment card, everyone receives the same set of thresholds. The fact that the referendum thresholds vary is what allows us to identify the location and scale of the conditional distribution of valuations.

observed *lower* bounds on the chosen intervals. The threshold assignments, for each Monte Carlo sample, were made as follows. We sorted (by size) all of the lower bounds for the *chosen* payment card intervals and tabulated how many times each bound occurred in our sample of 342 responses. We then generated 342 random integers on the range of 1 to 342 and assigned thresholds on this basis. For example, the lower bound of $15 was chosen for sorted observations 105 through 153. If the random integer generated for a given respondent happened to take on a value in this range, we assigned a referendum threshold (t_i) of $15 for that person. Next, if the lower bound of the interval that individual *actually* chose was $25 (for example), we would infer that this person would have responded "yes" ($I_i = 1$) to the question of whether they would be willing to pay $15.

The second distribution mimics the observed distribution of *upper* bounds on the chosen payment card intervals. The Monte Carlo assignment of thresholds "drawn from" this distribution was analogous to the procedure for the first distributional assumption.[10]

Two hundred Monte Carlo samples were generated in each case. Starting values for the maximum likelihood optimization algorithm were taken from the ML interval estimates for the "brief" model.[11] As for all the other ML models

---

[10] We initially assigned each of the twenty thresholds according to the outcomes from a uniform random integer generator (on the range of 1 to 20). This led to estimation problems, though, because by far the largest proportion of respondents implied valuations among the lower intervals. This resulted in a very large proportion of "no" responses in the Monte Carlo samples because the occurrence of high-valued thresholds was drastically disproportionate relative to the actual valuations in the sample. Some of the Monte Carlo simulations failed to converge because of this.

[11] The full model persisted in generating wildly implausible fitted values for some observations for three of the Monte Carlo samples, due to the magnitudes of the insignificant coefficients on the average catch rates. Deleting these variables made the fitted model more robust to the referendum threshold assignments, so we present results only for the "brief" model.

estimated in this paper, we use the package of FORTRAN subroutines called GQOPT. Rough optimization was achieved by the DFP procedure (using only first derivatives). Fine-tuning of the estimates (to a convergence tolerance of 10^{-10}) was accomplished by GRADX, a quadratic hill-climbing algorithm. The first column of Table VI repeats the ML interval estimates from Table II. Now, however, we include the asymptotic standard error estimates for each parameter in addition to the asymptotic t-test statistics. In conventional estimation, we use asymptotic standard errors for parameter estimates as a proxy for the degree of dispersion that we might expect in repeated sampling from the same population. Unfortunately, in most cases, we have only one sample at our disposal. In Monte Carlo experiments, however, we *generate* a large number of random samples. This enables us to actually calculate the dispersion in the point estimates across different samples. While the standard errors in the first column of Table VI are not directly comparable with the calculated standard errors (across 200 Monte Carlo experiments) reported in the second and third columns of the table, their differences give us a good idea about the extent of the loss in statistical efficiency when we move from a payment card question to a referendum question. Descriptive statistics for the results of the 200 Monte Carlo experiments, for the two threshold assignment schemes, make up the rest of Table VI.

The referendum slope estimates tend on average to overestimate the "true" ML interval slope estimates, although this could still be an artifact of the algorithm for the assignment of thresholds. We cannot state conclusively whether the apparent bias in point estimates under the referendum format is a robust result, although the two threshold assignment schemes we have considered are probably the most logical arbitrary choices.

The important lesson to be drawn from these Monte Carlo exercises is that the degree of variation in some of the point estimates is very high. In Table VI, we also provide "t-statistics" loosely constructed from the mean point estimate divided by the standard deviation (across the 200 samples). For all of the slope coefficients, a displacement of two standard deviations can change the sign of the point estimate. This suggests that a very wide range of point estimates (with vastly differing policy implications) can result from referendum-style data collection. Large samples will be necessary to minimize this "noise."

It is worthwhile to focus briefly on the range of possible implications that might be drawn across the different sets of randomly assigned thresholds in the Monte Carlo experiments. Tables VII(a) and VII(b) (using medians and means, respectively) detail means and standard deviations across the 200 different artificial samples of both the average derivatives of valuation with respect to each explanatory variable and the marginal mean fitted willingness-to-pay. Because using means as fitted values for WTP yields bigger numbers than using medians, the average derivatives summarized in Table VII(b) are considerably higher than those in Table VII(a).

In passing, we note that across these 200 samples, the average within-sample prediction success rate for responses to the fabricated yes/no WTP question were 77.30% and 76.59% for the "lower bound" and the "upper bound" models respectively. (The standard deviations of prediction success were 3.886% and 4.550%.) Regarding fitted valuations, the lower panel of Table VII(a) shows that the mean of the 200 fitted marginal mean WTP estimates (using medians) is $29.70 in the "lower bound" experiments, and $26.51 in the "upper bound" experiments. These average values, fortunately, are no further from the corresponding ML interval estimate ($31.22) than is the OLS midpoint method estimate ($26.03). When we use means to compute fitted WTP for each sample, as in Table VII(b), the average values are of course somewhat different.

9. Summary, Caveats, and Suggestions for Further Research

The primary finding of the econometric exercises reported thus far is that the implications of a contingent valuation study can be influenced to a considerable extent by both (a.) the valuation question format (payment card or referendum; payment card interval widths), and (b.) the estimation technique employed to fit the inverse Hicksian demand function. However, it is important to qualify these findings.

First, we can't know the actual values of the "true" underlying population parameters for the relationships attested to by the real data employed in this study--i.e. the true "data generating process." Therefore, we are at a loss to compare our "best" ML interval estimates (for the actual "payment card" elicitation method) with these true parameters to assess the sign or extent of the innate bias in the payment card format. Findings reported in Cameron (1987a) suggest that there will be some bias and that its sign will be indeterminate. Neither can we say whether the log-linear specification we are estimating is the "true" functional relationship between willingness to pay and the other variables. Evidence examined in Appendix 3 suggests that the log-linear specification may be too restrictive--a Box-Cox transformation employed with the ML interval method (which frees up an additional parameter) rejects the restriction imposed by the log-linear form. To minimize the complexity of comparing the different estimation methods and the different question formats, however, we have opted to adhere to the log-linear model and its relatively easier interpretation. One would certainly expect the sorts of discrepancies due to inappropriate estimation methods (or different interval specifications) to carry through for a variety of specifications of the relationship between WTP and the vector of x variables.

We have already emphasized that the method for assigning single thresholds to each respondent may contribute to the apparent bias in the point estimates. Nevertheless, in real referendum surveys, the thresholds are typically assigned randomly, with greater frequencies for the lower-valued thresholds. *Ex ante*, there does not seem to be a standard assignment algorithm to apply. Our arbitrary decisions are as plausible as any which would be made in practice. The bias in the referendum mean values of the Monte Carlo point estimates (relative to the ML interval estimates from the payment card data) is nevertheless fairly small. Certainly, the typical extent of the bias evidenced by the artificial referendum data is not much worse than that generated by reliance on the inappropriate OLS midpoint method with payment card data. This is somewhat heartening. Still, more work remains to be done on this topic. The current paper utilizes real data--a *purely* Monte Carlo study may be necessary to answer some of the questions which remain unresolved because we cannot know the true data generating process for our sample.

This study generates information which will be of use to the designers of contingent valuation surveys and those responsible for deciding upon sample sizes. Users of contingent valuation survey data need a careful study of the relative rates of questionnaire completion across different "vehicles" before we can assess which format will be appropriate in specific situations. If respondents can "take in" the entire payment card at a glance and decide easily which interval brackets their true value (without bias), then our results would seem to point strongly in favor of the payment card vehicle, given the substantial loss of information with the referendum approach.

However, we cannot know the extent to which the presentation of the information in the payment card will bias the results, or whether the intervals are too finely specified (i.e. intervals which are too narrow present the respondent with the same dilemma as do "open-ended" contingent valuation questions).

In sum, we have gained some valuable insights into the perils for contingent valuation studies of using inappropriate estimation methods (the OLS midpoint technique) with payment card data. Section 7, especially, focused on the dependence of interval estimates upon the number of intervals specified. We have also established that while referendum data collection methods *may* yield widely varying point estimates in samples as small as this one, it seems to be the case that *on average*, the point estimates are only slightly biased. This is a good argument for collecting the largest possible sample if the referendum question format is to be used. Overall, then, the policy implications of a contingent valuation study have the potential to be highly sensitive to the design of the questionnaire.

Table I

Descriptive Statistics
(n = 342)

Variable	Description	Unweighted Mean (std. dev.)	Weighted[a] Mean (std. dev.)
MIDPT	midpoint of interval	48.66 (80.49)	57.98 (132.96)
log(MIDPT)	log (midpoint of interval)	3.165 (1.273)	3.115 (1.371)
TRPS	# salmon and striped bass fishing trips in past 12 mo.	5.591 (6.835)	4.416 (5.449)
log(INC)	log($'000 hhld income using midpoint of reported interval)	3.647 (0.6447)	3.602 (0.6544)
STRIP	= 1 if all trips were salmon fishing trips	0.4298	0.4188
BTRIP	= 1 if all trips were striped bass fishing trips	0.3743	0.3338
ABIL2	= 1 if intermediate fishing ability ("3" on 1-5)	0.4035	0.4272
ABIL3	= 1 if advanced fishing ability ("4", "5" on 1-5)	0.3363	0.2813
OWNBOAT	= 1 if respondent owns a boat	0.4035	0.3317
[b]C1SAL	catch/trip of salmon on excl. salmon trips	0.7544 (1.191)	0.7341 (1.046)
C2SB	catch/trip of bass on excl. bass trips	0.4771 (1.080)	0.4446 (0.9844)
C3SB	catch/trip of bass on salmon/bass trips	0.06433 (0.4966)	0.06861 (0.5023)
C3SAL	catch/trip of salmon on salmon/bass trips	0.06775 (0.3538)	0.08253 (0.3767)

[a] weights were derived from cross-tabulations of home county by frequency of trip for both the original sample and the estimation sample. Deletion criteria were: number of trips unknown; income unknown; ability unknown; boat ownership unknown; total number of saltwater angling trips in the last 12 months unknown.

[b] simple averages; values can be zero of no trips of the specified type were taken. Since data are retrospective over last three trips, we cannot simply use STRIP and BTRIP to compute actual per-trip catch--some anglers will have reported mixed trip types.

Table II

Comparison of OLS Midpoint Estimates and ML Interval Estimates

Variable	OLS Midpoint Estimates		ML Interval Estimates	
constant	2.180 (4.812)[a]	2.251 (4.434)	2.452 (7.626)	2.530 (8.736)[b]
TRPS	0.02902 (1.971)	0.02908 (1.989)	0.02951 (2.864)	0.02862 (2.861)
log(INC)	0.3591 (3.275)	0.3425 (3.196)	0.2759 (3.531)	0.2531 (3.306)
STRIP	-0.5248 (-2.598)	-0.4559 (-2.465)	-0.1199 (-0.8484)	-
BTRIP	-0.6984 (-3.135)	-0.5816 (-2.985)	-0.5128 (-3.252)	-0.4495 (-3.631)
ABIL2	0.2155 (1.212)	-	0.1446 (1.132)	-
ABIL3	0.7120 (3.447)	0.5609 (3.301)	0.3622 (2.481)	0.2781 (2.308)
OWNBOAT	-0.4757 (-3.0462)	-0.4412 (-2.871)	-0.1360 (-1.180)	-
C1SAL	-0.1986 (-2.3734)	-0.1690 (-2.111)	-0.1372 (-2.285)	-0.1497 (-2.688)
C2SB	-0.004381 (-0.05366)	-	0.01755 (0.3105)	-
C3SB	-0.06630 (-0.4586)	-	-0.05635 (-0.5615)	-
C3SAL	-0.1910 (-0.9155)	-	0.07988 (0.5409)	-
σ	0.8947	1.286	0.8679 (23.22)	0.8724 (23.25)
max log L	-1631.60	-1632.77	-653.47	-655.62

[a] t-ratios from OLS regression output (SHAZAM)

[b] asymptotic t-test statistics from GQOPT output (GRADX)

Table III(a)

Contrasting the Implications of the Log-linear Model Estimates by OLS and by the ML Interval Method: MEDIANS

	OLS Estimates		ML Method	
Variable	Full Model	Brief Model	Full Model	Brief Model
mean ∂WTP/∂x (with standard deviations):				
TRPS	\$ 0.7615[b] (0.5142)	\$ 0.7571 (0.4909)	\$ 0.9250 (0.3718)	\$ 0.8936 (0.3271)
log(INC)[a]	9.423 (6.363)	8.916 (5.781)	8.647 (3.475)	7.901 (2.892)
STRIP	-13.77 (9.298)	-11.87 (7.695)	-3.758 (1.510)	-
BTRIP	-18.33 (12.38)	-15.14 (9.816)	-16.07 (6.459)	-14.03 (5.137)
ABIL2	5.656 (3.819)	-	4.533 (1.822)	-
ABIL3	18.68 (12.62)	14.60 (9.467)	11.35 (4.562)	8.680 (3.178)
OWNBOAT	-12.48 (8.428)	-11.49 (7.447)	-4.264 (1.714)	-
C1SAL	-5.210 (3.518)	-4.400 (2.853)	-4.299 (1.728)	-4.674 (1.711)
C2SB	-0.1150 (0.07763)	-	0.5501 (0.2211)	-
C3SB	-1.740 (1.175)	-	-1.766 (0.7097)	-
C3SAL	-5.012 (3.385)	-	2.503 (1.006)	-
mean WTP (with standard deviation):				
	\$ 26.24 (17.72)	\$ 26.03 (16.88)	\$ 31.39 (12.60)	\$ 31.22 (11.47)

[a] derivatives with respect to log(INC), not INC itself

[b] point estimate of parameter times exponentiated fitted value from log-linear specification, weighted average across all observations

Table III(b)

Contrasting the Implications of the Log-linear Model Estimates by OLS and by the ML Interval Method: MEANS

	OLS Estimates		ML Method	
Variable	Full Model	Brief Model	Full Model	Brief Model
mean ∂WTP/∂x (with standard deviation):				
TRPS	\$ 1.136[b]	\$ 1.731	\$ 1.348	\$ 1.307
	(0.7673)	(1.122)	(0.5418)	(0.4786)
log(INC)[a]	14.06	20.38	12.60	11.56
	(9.495)	(13.22)	(5.064)	(4.231)
STRIP	-20.55	-27.14	-5.477	-
	(13.87)	(17.59)	(2.201)	
BTRIP	-27.35	-34.61	-23.42	-20.53
	(18.47)	(22.44)	(9.413)	(7.516)
ABIL2	8.440	-	6.606	-
	(5.699)		(2.655)	
ABIL3	27.87	33.38	16.54	12.70
	(18.83)	(21.64)	(6.648)	(4.650)
OWNBOAT	-18.62	-26.27	-6.214	-
	(12.58)	(17.02)	(2.498)	
C1SAL	-7.774	-10.06	-6.265	-6.838
	(5.250)	(6.523)	(2.518)	(2.503)
C2SB	-0.1716	-	0.8017	-
	(0.1158)		(0.3222)	
C3SB	-2.596	-	-2.574	-
	(1.753)		(1.034)	
C3SAL	-7.479	-	3.648	-
	(5.051)		(1.466)	
mean WTP (with standard deviation):				
	\$ 39.16	\$ 59.51	\$ 45.75	\$ 45.68
	(26.44)	(38.59)	(18.36)	(16.78)

[a] derivatives with respect to log(INC), not INC itself

[b] point estimate of parameter times exponentiated fitted value from log-linear specification, times $\exp(\sigma^2/2)$; weighted average across all observations

Table IV(a)

Comparison of ML Interval Estimates for Original Intervals versus Coarser intervals ("Brief" Model Only)

Variable	original[a] intervals	double[b] intervals	triple[c] intervals	quadruple[d] intervals
point estimates (asymptotic t-statistics):				
constant	2.530 (8.736)	2.740 (9.392)	3.695 (14.81)	3.808 (16.81)
TRPS	0.02862 (2.861)	0.02933 (2.943)	0.02101 (2.544)	0.01183 (1.788)
log(INC)	0.2531 (3.306)	0.2165 (2.807)	0.08494 (1.319)	0.1157 (2.023)
BTRIP	-0.4495 (-3.631)	-0.4332 (-3.471)	-0.5321 (-4.936)	-0.2982 (-3.117)
ABIL3	0.2781 (2.308)	0.2385 (1.975)	0.1413 (1.388)	0.2518 (2.818)
C1SAL	-0.1497 (-2.688)	-0.1446 (-2.607)	-0.1780 (-3.748)	-0.1966 (-4.225)
σ	0.8724 (23.25)	0.8504 (22.32)	0.6268 (17.93)	0.3889 (9.647)

[a] reproduced to facilitate comparison. Intervals: $0-5, 5-10, 10-15, 15-20, 20-25, 25-50, 50-75, 75-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-550, 550-600, 600-750, >750.

[b] Intervals: $0-10, 10-20, 20-50, 50-100, 100-200, 200-300, 300-400, 400-500, 500-600, >600.

[c] Intervals: $0-15, 15-50, 50-150, 150-300, 300-450, 450-600, >600

[d] Intervals: $0-20, 20-100, 100-300, 300-500, >500.

Table IV(b)

Contrasting the Implications of ML estimates for Original Intervals versus Coarser intervals ("Brief" Model Only): MEDIANS

Variable	original[a] intervals	double intervals	triple intervals	quadruple intervals
mean ∂WTP/∂x (with standard deviations):				
TRPS	$ 0.8936 (0.3271)	$ 0.9842 (0.3377)	$ 0.9947 (0.2672)	$ 0.7375 (0.1793)
log(INC)	7.901 (2.892)	7.266 (2.493)	4.022 (1.080)	7.213 (1.754)
BTRIP	-14.03 (5.137)	-14.54 (4.988)	-25.20 (6.768)	-18.59 (4.520)
ABIL3	8.680 (3.178)	8.005 (2.746)	6.691 (1.798)	15.69 (3.816)
C1SAL	-4.674 (1.711)	-4.851 (1.664)	-8.428 (2.264)	-12.75 (2.979)
mean WTP (with standard deviations):				
	$ 31.22 (12.60)	$ 33.56 (11.51)	$ 47.35 (12.72)	$ 62.32 (15.16)

[a] intervals as in Table IV(a)

Table IV(c)

Contrasting the Implications of ML estimates for Original Intervals versus Coarser intervals ("Brief" Model Only): MEANS

Variable	original[a] intervals	double intervals	triple intervals	quadruple intervals
mean ∂WTP/∂x (with standard deviations):				
TRPS	\$ 1.307 (0.4786)	\$ 1.413 (0.4848)	\$ 1.211 (0.3252)	\$ 0.7954 (0.1934)
log(INC)	11.56 (4.231)	10.43 (3.579)	4.895 (1.314)	7.780 (1.892)
BTRIP	-20.53 (7.516)	-20.87 (7.161)	-30.67 (8.237)	-20.05 (4.875)
ABIL3	12.70 (4.650)	11.49 (3.942)	8.143 (2.188)	16.92 (4.116)
C1SAL	-6.838 (2.503)	-6.964 (2.389)	-10.26 (2.755)	-13.75 (3.213)
mean WTP (with standard deviations):				
	\$ 45.68 (18.43)	\$ 48.18 (16.52)	\$ 57.63 (15.48)	\$ 67.22 (16.35)

[a] intervals as in Table IV(a)

Table V(a)

Comparison of ML Interval Estimates for Original Intervals versus Fewer Upper Intervals("Brief" Model Only)

Variable	original[a] intervals (> 750)	highest interval (> 300)[b]	highest interval (> 200)[c]	highest interval (> 100)[d]
point estimates (asymptotic t-statistics):				
constant	2.530 (8.736)	2.570 (9.010)	2.844 (10.84)	2.769 (11.92)
TRPS	0.02862 (2.861)	0.03027 (3.073)	0.0295 (2.283)	0.01702 (1.913)
log(INC)	0.2531 (3.306)	0.2420 (3.209)	0.1357 (1.935)	0.1264 (2.037)
BTRIP	-0.4495 (-3.631)	-0.4703 (-3.858)	-0.2630 (-2.277)	-0.1939 (-1.879)
ABIL3	0.2781 (2.308)	0.2487 (2.095)	0.1046 (0.9393)	0.08299 (0.8462)
C1SAL	-0.1497 (-2.688)	-0.1486 (-2.712)	-0.07144 (-1.376)	-0.02003 (-0.4424)
σ	0.8724 (23.25)	0.8580 (23.19)	0.7820 (22.51)	0.6814 (21.52)

[a] reproduced to facilitate comparison. Intervals: $0-5, 5-10, 10-15, 15-20, 20-25, 25-50, 50-75, 75-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-550, 550-600, 600-750, >750. Highest interval has one observation.

[b] Highest interval now contains 2 observations.

[c] Highest interval now contains 12 observations.

[d] Highest interval now contains 46 observations.

Table V(b)

Contrasting the Implications for ML Interval Estimates for Original Intervals versus Fewer Upper Intervals ("Brief" Model Only): MEDIANS

Variable	original[a] intervals (> 750)	highest interval (> 300)	highest interval (> 200)	highest interval (> 100)
mean ∂WTP/∂x (standard deviation):				
TRPS	$ 0.8936 (0.3271)	$ 0.9375 (3.430)	$ 0.5864 (0.1166)	$ 0.4411 (0.07153)
log(INC)	7.901 (2.892)	7.495 (2.740)	3.799 (0.7553)	3.274 (0.5309)
BTRIP	-14.03 (5.137)	-14.57 (5.330)	-7.363 (1.464)	-5.025 (0.8148)
ABIL3	8.680 (3.178)	7.703 (2.818)	2.929 (0.5823)	2.150 (0.3487)
C1SAL	-4.674 (1.711)	-4.601 (1.683)	-2.000 (0.3976)	-0.5192 (0.08420)
mean WTP (with standard deviations):				
	$ 31.22 (12.60)	$ 30.97 (11.33)	$ 27.99 (5.565)	$ 25.91 (4.202)

[a] intervals as in Table V(a).

Table V(c)

Contrasting the Implications for ML Interval Estimates for Original Intervals versus Fewer Upper Intervals ("Brief" Model Only): MEANS

Variable	original[a] intervals (> 750)	highest interval (> 300)	highest interval (> 200)	highest interval (> 100)
mean ∂WTP/∂x (standard deviation):				
TRPS	\$ 1.307 (0.4786)	\$ 1.355 (4.956)	\$ 0.7961 (0.1583)	\$ 0.5564 (0.09022)
log(INC)	11.56 (4.231)	10.83 (3.959)	5.158 (1.0253)	4.130 (0.6696)
BTRIP	-20.53 (7.516)	-21.05 (7.702)	-9.996 (1.988)	-6.338 (1.028)
ABIL3	12.70 (4.650)	11.13 (4.072)	3.977 (0.7906)	2.712 (0.4398)
C1SAL	-6.838 (2.503)	-6.648 (2.432)	-2.715 (0.5398)	-0.6549 (0.1062)
mean WTP (with standard deviations):				
	\$ 45.68 (18.43)	\$ 44.75 (16.37)	\$ 38.00 (7.555)	\$ 32.68 (5.300)

[a] intervals as in Table V(a).

Table VI

ML Interval Estimates versus Simulated Referendum Data
(200 Monte Carlo Samples; Referendum Threshold Frequencies
Matching Payment Card Lower Bound Frequencies)

Variable	ML Interval[a]	Monte Carlo Referendum "lower bound"	Monte Carlo Referendum "upper bound"
constant	2.530	2.269	1.969
	(0.2896)[b]	(0.7793)[c]	(0.7671)
	(8.736)	(2.911)[d]	(2.567)
TRPS	0.02862	0.03569	0.03144
	(0.0100)	(0.02087)	(0.01836)
	(2.861)	(1.711)	(1.712)
log(INC)	0.2531	0.2652	0.3246
	(0.07656)	(0.1950)	(0.1939)
	(3.306)	(1.360)	(1.674)
BTRIP	-0.4495	-0.5217	-0.4763
	(0.1238)	(0.3184)	(0.2734)
	(-3.631)	(-1.639)	(-1.742)
ABIL3	0.2781	0.4211	0.4791
	(0.1205)	(0.2978)	(0.2602)
	(2.308)	(1.414)	(1.842)
C1SAL	-0.1497	-0.2008	-0.2454
	(0.05569)	(0.1394)	(0.1412)
	(-2.688)	(-1.441)	(-1.737)
σ	0.8724	1.195	1.231
	(0.03752)	(0.2666)	(0.2396)
	(23.25)	(4.481)	(5.135)
max log L[e]	-655.62	-150.70	-160.90
		(14.88)	(17.17)

[a] reproduced to facilitate comparison

[b] asymptotic standard errors and asymptotic t-test statistics (GQOPT output, GRADX routine)

[c] standard deviation of point estimates across 200 Monte Carlo simulations

[d] analogy to t-value computed using standard deviation of Monte Carlo point estimates across 200 samples.

[e] maximized values of the log-likelihood functions are of course not comparable, since the models are non-nested. Conceptually, it is of course easier to predict residence of an observation in the wider "intervals" defined by the referendum format.

Table VII(a)

Distribution of Implications Drawn from Referendum Data

(Means and standard deviations, 200 Monte Carlo samples): MEDIANS

	"lower bound"	"upper bound"
Mean ∂WTP/∂x (with standard deviations, across 200 samples):		
TRPS	$ 1.104 (0.8077)	$ 0.8583 (0.5934)
log(INC)	8.038 (6.642)	8.770 (5.650)
BTRIP	-16.30 (12.38)	-13.05 (8.474)
ABIL3	12.87 (9.965)	13.00 (8.055)
C1SAL	-6.259 (4.994)	-6.671 (4.397)
Mean WTP (with standard deviations, across 200 samples):		
	$ 29.70 (4.968)	$ 26.51 (3.805)

Table VII(b)

Distribution of Implications Drawn from Referendum Data

(Means and standard deviations, 200 Monte Carlo samples): MEANS

	"lower bound"	"upper bound"
Mean ∂WTP/∂x (with standard deviations, across 200 samples):		
TRPS	$ 2.946 (4.624)	$ 2.200 (2.433)
log(INC)	19.38 (30.19)	20.98 (18.49)
BTRIP	-42.80 (64.40)	-32.32 (30.22)
ABIL3	34.22 (53.76)	33.10 (32.34)
C1SAL	-17.14 (26.52)	-17.32 (18.18)
Mean WTP (with standard deviations, across 200 samples):		
	$ 70.66 (55.04)	$ 62.60 (31.96)

REFERENCES

Burridge, J. (1982) "A Note on Maximum Likelihood Estimation for Regression Models Using Grouped Data," *Journal of the Royal Statistical Society*, Series B, 43, 41-45.

Cameron, T.A. (1987a) "The Impact of Grouping Coarseness in Alternative Grouped-Data Regression Models," forthcoming, *Journal of Econometrics* (Annals).

Cameron, T.A. (1987b) "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 (1986) "The Determinants of Value for a Recreational Fishing Day: Estimates from a Contingent Valuation Survey," Discussion Paper #405, Department of Economics, University of California, Los Angeles.

Cameron, T.A., and M.D. James (1987) "Efficient Estimation Methods for Use with 'Closed-Ended' Contingent Valuation Survey Data," *Review of Economics and Statistics*, 69, no. 2, (May) pp. 269-276.

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

Goldfeld, S. and R.E. Quandt, *GQOPT*: A Package for Numerical Optimization of Functions, Department of Economics, Princeton University, Princeton, N.J. 08544.

Greene, W.H. (October, 1986) *LIMDEP* (trademark) User's Manual

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

Hasselblad, V., Stead, A.G., and Galke, W. (1980) "Analysis of Coarsely Grouped Data from the Lognormal Distribution," *Journal of the American Statistical Association*, December, 75, No. 372, pp. 771-778.

Hastings, N.A.J., and J.B. Peacock (1974) *Statistical Distributions*, New York: Wiley.

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

Stewart, M.B. (1983) "On Least Squares Estimation When the Dependent Variable is Grouped," *Review of Economic Studies*, 50, 737-53.

Thompson, Cynthia J., and Daniel D. Huppert (1987) "Results of the Bay Area Sportfish Economic Study (Bases)," U.S. Department of Commerce, NOAA Technical Memorandum, National Marine Fisheries Service (July).

APPENDIX 1: DERIVATIVES FOR THE "PAYMENT CARD" MODEL

To simplify the formulas for the gradient vector and the Hessian matrix, it is useful to define the following abbreviations:

$$P_i = \Phi(z_{ui}) - \Phi(z_{li})$$

$$p_i = \phi(z_{ui}) - \phi(z_{li})$$

$$Q_i = z_{ui}\,\phi(z_{ui}) - z_{li}\,\phi(z_{li})$$

$$q_i = z_{li}^{\;2}\,\phi(z_{li}) - z_{ui}^{\;2}\,\phi(z_{ui})$$

$$R_i = z_{li}^{\;3}\,\phi(z_{li}) - z_{ui}^{\;3}\,\phi(z_{ui})$$

$$\rho_i = p_i/P_i$$

$$\omega_i = Q_i/P_i$$

The elements of the gradient vector are given by:

$$\partial \log L/\partial \beta_r = \sum - \rho_i x_{ir}/\sigma \qquad r = 1,\ldots,p$$

$$\partial \log L/\partial \sigma = \sum - \omega_i/\sigma$$

and the elements of the Hessian are:

$$\partial^2 \log L/\partial \beta_r \partial \beta_s = \sum \left[\rho_i(1-\rho_i)\right] x_{ir} x_{is}/\sigma^2 \qquad r,s = 1,\ldots,p$$

$$\partial^2 \log L/\partial \beta_r \partial \sigma = \sum \left[\rho_i + q_i/P_i - \rho_i\omega_i\right] x_{ir}/\sigma^2 \qquad r = 1,\ldots,p$$

$$\partial^2 \log L/\partial \sigma^2 = \sum \left[2\,\omega_i + R_i/P_i - \omega_i^{\;2}\right]/\sigma^2$$

APPENDIX 2: DERIVATIVES FOR THE "REFERENDUM" MODEL

Using the notation established in the text, we first define the following simplifying abbreviations (z denotes the standard normal random variable in this appendix):

$$z_i = (t_i - x_i'\beta)/\sigma$$

$$\Phi_i = \Phi(z_i) \qquad \phi_i = \phi(z_i) \qquad \phi'_i = \phi'(z_i) = -z_i\,\phi(z_i)$$

$$R_i = x_{ir}x_{is}\,\phi'_i \qquad S_i = x_{ir}x_{is}\,\phi^2_i$$

$$T_i = x_{ir}z_i\,\phi'_i \qquad U_i = x_{ir}z_i\,\phi^2_i$$

$$V_i = z^2_i\,\phi'_i \qquad W_i = z^2_i\,\phi^2_i$$

The gradient vector for this model is then given by:

$$\partial log\ L/\partial\beta_r = (1/\sigma) \sum \{ [I_i - (1 - \Phi_i)]\ x_{ir}\ \phi_i / [\Phi_i\ (1 - \Phi_i)]\} \qquad r = 1,\ldots,p$$

$$\partial log\ L/\partial\sigma = (1/\sigma) \sum \{ [I_i - (1 - \Phi_i)]\ z_i\ \phi_i / [\Phi_i\ (1 - \Phi_i)]\}$$

The elements of the Hessian matrix can be simplified if we define the function:

$$G(P,Q) = \sum \left[\frac{I_i(P_i[\Phi_i - 1] - Q_i)}{[\Phi_i - 1]^2} + \frac{(1 - I_i)(P_i\Phi_i - Q_i)}{\Phi^2_i} \right]$$

Then:

$$\partial^2 logL/\partial\beta_r\partial\beta_s = (1/\sigma)\ G(R,S) \qquad r,s = 1,\ldots,p$$

$$\partial^2 logL/\partial\beta_r\partial\sigma = (-1/\sigma)\ \partial logL/\partial\beta_r + (1/\sigma^2)\ G(T,U) \qquad r = 1,\ldots,p$$

$$\partial^2 logL/\partial\sigma^2 = (-1/\sigma)\ \partial logL/\partial\sigma + (1/\sigma^2)\ G(V,W)$$

Use of these analytic derivatives instead of numerical approximations can reduce computational costs considerably.

APPENDIX 3: ML INTERVAL RESULTS WITH BOX-COX TRANSFORMATION

Although a log-linear specification is adopted in the body of this paper (since our emphasis is primarily upon the different estimates achieved by the payment card versus the referendum value elicitation methods), we did explore alternative transformations of the underlying continuous dependent variable. Specifically, we have experimented with a Box-Cox transformation: $(Y_i^{\lambda} - 1)/\lambda = x_i'\beta$ (with the transformation parameter λ estimated simultaneously with the β vector and σ by ML). The results for this specification, by the ML interval method, are displayed in Table A.3.I

Frequently, the estimated Box-Cox transformation parameter λ will lie between zero and unity. Here, however, it is a small negative number, on the order of -0.23. (A plot of the Box-Cox transformation with this parameter value is more sharply curved than a logarithmic transformation; larger values are therefore shrunk even more severely than would be suggested by a log transformation.) While the λ is indeed statistically significantly different from zero (which would imply a log transformation), we elect to utilize the log transformation primarily because of its simplicity. Of course, since the Box-Cox model has an additional parameter free to be determined by the evidence in the sample, these estimates reflect closer conformity to the relationship between the variables implied by the data. These estimates can therefore be considered "superior" to any of those appearing in the body of the paper, as far as *ad hoc* models are concerned.

In the body of the paper, we took care to distinguish between the fitted conditional *median* values for WTP and the fitted conditional *mean* WTP estimates. In the simple log-linear specification, the latter could be obtained from the former by transforming by $\exp(\sigma^2/2)$. In the Box-Cox example, the correction is not this simple. Specifically, the fitted value of

$x_i'\beta$ gives the mean of the transformed variable, $Y^{(\lambda)} = (Y^\lambda - 1)/\lambda$. The *mean* of Y itself will be given by the expectation of a function, $[\lambda\ Y^{(\lambda)} + 1]^{(1/\lambda)}$, of the $N(x_i'\beta,\ \sigma^2)$ random variable, $Y^{(\lambda)}$. Since this integral seems prohibitively difficult to evaluate, we will report only the "de-transformed" value of the mean of $Y^{(\lambda)}$.

Recall that the marginal mean of fitted WTP values (medians) in the "brief" log-linear ML interval model was $31.22 (with a standard deviation of $11.47). In the "brief" Box-Cox model, the corresponding marginal mean is $28.78 (with a standard deviation of 9.908). These differences are relatively modest. From a policy-making point of view, they may or may not be "significant."

Table A.3.I

ML Interval Estimates with
Box-Cox Transformed Dependent Variable

Variable	Interval Estimates (asy.t-ratio)	Mean Implied Derivatives (std.dev.)	Interval Estimates (asy.t-ratio)	Mean Implied Derivatives (std.dev.)
constant	1.898 (9.473)[a]	-	1.936 (9.931)	-
TRPS	0.01230 (2.255)	0.7933 (0.3853)[b]	0.01190 (2.287)	0.7718 (0.3484)
log(INC)	0.1152 (2.580)	7.428 (3.607)	0.1030 (2.462)	6.684 (3.017)
STRIP	-0.03511 (-0.5680)	-2.264 (1.100)	-	-
BTRIP	-0.2096 (-2.396)	-13.52 (6.565)	-0.1770 (-2.489)	-11.48 (5.182)
ABIL2	0.06689 (1.170)	4.314 (2.095)	-	-
ABIL3	0.1493 (2.020)	9.631 (4.677)	0.1086 (1.848)	7.046 (3.180)
OWNBOAT	-0.04757 (-0.9217)	-3.068 (1.490)	-	-
C1SAL	-0.05436 (-1.826)	-3.506 (1.703)	-0.05517 (-1.937)	-3.580 (1.616)
C2SB	0.01399 (0.5636)	0.9022 (0.4382)	-	-
C3SB	-0.01798 (-0.4151)	-1.160 (0.5631)	-	-
C3SAL	0.02672 (0.4199)	1.723 (0.8368)	-	-
σ	0.3720 (3.932)	-	0.3677 (3.949)	-
λ	-0.2329 (-3.306)	-	-0.2374 (-3.383)	-
max log L	-647.69		-649.53	

[a] asymptotic t-test statistics from GQOPT

[b] standard deviation of fitted derivatives over sample