Preliminary Version 02/11/87

VALUING PUBLIC GOODS USING REFERENDUM DATA: ESTIMATION ASSUMING A LOGISTIC ERROR DISTRIBUTION

by

Trudy Ann Cameron*
Department of Economics, UCLA

Discussion Paper # 430
February 1987

* The data used in this study were collected under a project sponsored by the Department of Fisheries and Oceans, Vancouver, Canada, supervised by Michelle D. James, my co-author on earlier related papers. The comments of four anonymous referees on these earlier papers have also contributed to the evolution of the present paper. E. Jane Murdoch has offered helpful suggestions.

Valuing Public Goods Using Referendum Data: Estimation Assuming a Logistic Error Distribution[1]

by

Trudy Ann Cameron
Department of Economics, UCLA

Proposed running head:

Logistic Models for Referendum Data

Direct correspondence to:

Professor T. A. Cameron
Department of Economics
University of California
405 Hilgard Avenue
Los Angeles, CA 90024
(213)825-3925
(213)825-1011

ABSTRACT

In an earlier study, we developed a reparameterization of a conventional maximum likelihood *probit* model which allows estimation of regression-like coefficients with variable-threshold "referendum" data. The procedure, however, requires repeated evaluation of the integral of the standard normal density function. Since this can be expensive, the present paper describes a logistic analog, where the requisite integrals are simple closed forms. Additionally, since earlier researchers have based their analyses on conventional maximum likelihood logit estimates, we can show how their point estimates can be transformed very simply to yield precise formulas for their underlying inverse demand functions.

Valuing Public Goods Using Referendum Data: Estimation Assuming a Logistic Error Distribution

I. INTRODUCTION

"Closed-ended" contingent valuation (CECV) surveys (sometimes called "referendum" surveys), have become very popular as a technique for eliciting the value of public goods or non-market resources. For comprehensive assessments of these survey instruments, see either Cummings, Brookshire, and Schulze [6], or Mitchell and Carson [11]. Briefly, the procedure involves first establishing the attributes of the public good or the resource, and then asking the respondent whether or not they would pay or accept a single specific sum for access. The arbitrarily assigned sums are varied across respondents. This questioning strategy is attractive because it generates a scenario similar to that encountered by consumers in their usual market transactions. A hypothetical price is stated and the respondent merely decides whether to "take it or leave it." This is less stressful for the respondent than requiring that a specific value be named, and circumvents much of the potential for strategic response bias. The challenge for estimation arises because the respondent's true valuation is an unobserved random variable. We can only infer its magnitude through an indicator variable that tells us whether this underlying value is greater or less than the offered "threshold" value.

In two earlier papers (Cameron and James [4], and Cameron and James [3]), we designed a new maximum likelihood procedure to utilize data of this type. The first paper emphasizes the methodology and convenient approximations; the second paper is an in-depth study concerning the valuation of a recreational fishery. In both of these papers, we maintain distributional hypotheses consistent with those underlying ordinary least squares (OLS) regression models: i.e. that the population error terms have the familiar normal density function. The resulting parameters can therefore be interpreted in the same manner as those resulting from OLS regression.

Earlier studies utilizing closed-ended contingent valuation data have typically employed conventional maximum likelihood *logit* models to estimate parameters which can be combined with the data to predict the log-odds for the respondent's "yes" or "no" response to the offered amount. These studies adhere to simply binary choice formulations, with the offered value treated an an "explanatory" variable in the discrete choice model. Once these parameters are estimated, an estimate of mean valuation across the sample (evaluated at the means of the other explanatory variables) is calculated by determining the area under the fitted logit probability curve.[2] The limits of integration are from $0 to the highest offered amount. Likewise, to determine the effect on mean valuation of a one-unit change in some other explanatory variable, that particular variable is permuted by one unit for all observations, and the area under the recalculated probability curve is computed.

For policy, it is important to be able to assess the sensitivity of a resource's value to changes in the levels of factors which affect it. (E.g. if pollution levels affect the value of a recreational area, and the government desires to perform a cost-benefit analysis to determine the advisability of an expensive clean-up program, it is essential to know by how much a given decrease in the pollution level will affect the value of the area to particular types of individuals, as well as the aggregate social value of the area.) The difficulty with the earlier discrete-choice logit methods is that while approximate point estimates of these sensitivities can be deduced from the estimated model, it seems to be prohibitively difficult to assign standard errors for these estimates.

Our reparameterized probit model resolves this problem, and in our other papers, we have recommended this technique (with its underlying normality assumption) primarily on the basis of its correspondence to OLS models. Given that most practitioners and many policy-makers have some degree of familiarity with the interpretation of OLS models, the normality assumption seems particularly attractive. However, the normality assumption in discrete-choice-based models has one shortcoming which, incidentally, accounts for the emergence of logit models in the first place: it is expensive.

Discrete choice models require the computation of cumulative densities. For the normal distribution, the cumulative distribution is not of closed form--it is an integral which must be computed numerically. While there are accurate subroutines for computing these integrals, they must be computed a very large number of times during each iteration of a maximum likelihood optimization procedure. This makes computation of discrete choice models based on the normal distribution potentially very expensive (or at least time-consuming).

For logit models, however, which are based on the logistic distribution, the cumulative density does have a closed form. Its value is simply a ratio of exponentiated quantities which can be computed relatively cheaply and quickly. Since the shapes of the standard normal and the standard logistic distributions are almost identical (except that the latter is slightly thicker in the tails), the logistic-based model provides a very convenient and typically very accurate approximation to the normal-based model (which would otherwise be preferred on criteria for statistical inference).

As a second rationale for developing a logit-based CECV technique, it would be interesting to be able go back to published studies using the conventional logit model to disentangle their estimated coefficients and to uncover the underlying implicit inverse demand functions employed by the authors.

In section II, we describe the likelihood function for use with CECV data under the logistic assumption. In section III, we examine a highly simplified empirical example using a subset of the data from our previous study. We then re-examine some earlier published studies in section IV, suggesting that a "random utility maximization" approach may be unnecessary with CECV data.

II. LOG-LIKELIHOOD FUNCTION FOR CECV DATA WITH LOGISTIC ERRORS

This section is analogous to the discussion for a model with normal errors given in Cameron and James [5]. Assume that the unobserved continuous dependent variable is the respondent's true willingness-to-pay (*WTP*)[3] for the resource or public good, Y_i. If we assume that the underlying distribution of Y_i, conditional on a vector of explanatory variables, x_i, has a logistic (rather than a normal) distribution, with a mean of $x_i'\beta$, then maximum likelihood techniques are still appropriate.

In the standard binary logit model, we would assume that:

$$Y_i = x_i'\beta + u_i \tag{1}$$

where Y_i is unobserved, but is manifested through the discrete indicator variable, y_i, such that:

$$y_i = 1 \text{ if } Y_i > 0 \tag{2}$$
$$= 0 \text{ otherwise.}$$

If we assume that u_i is distributed according to a logistic distribution with mean 0 and standard deviation b (and with alternative parameter $k = b\sqrt{3}/\pi$, see Hastings and Peacock, 1975), then

$$\Pr(y_i = 1) = \Pr(Y_i > 0) = \Pr(u_i > -x_i'\beta) \tag{3}$$
$$= \Pr(u_i/k > -x_i'\beta/k)$$
$$= 1 - \Pr(\eta_i < -x_i'\gamma),$$

where $\gamma = \beta/k$ and we use η to signify the standard logistic random variable with mean 0 and standard deviation $b = \pi/\sqrt{3}$. The formula for the cumulative density up to z for the standard logistic distribution is

$$1 - \{1 + \exp[z]\}^{-1}. \tag{4}$$

Therefore the log-likelihood function can be written as:

$$\log L = \Sigma - y_i \log\{1 + \exp[-x_i'\gamma]\} + (1 - y_i) \log\{\exp[-x_i'\gamma]/(1 + \exp[-x_i'\gamma])\}. \tag{5}$$

Simplification[4] yields:

$$\log L = \Sigma (1 - y_i)(-x_i'\gamma) - \log[1 + \exp(-x_i'\gamma)]. \tag{6}$$

Note that it is not possible in this model to estimate β separately, since it always appears as β/k. The model must therefore be evaluated in terms of its estimated probabilities, since the underlying valuation function, $x_i'\beta$, cannot be recovered.

With CECV data, however, each individual is confronted with a threshold value, t_i, and by his (yes/no) response, we conclude that his true *WTP* is either greater than or less than t_i. Therefore, the conventional logit model can be modified. As before, we can assume a valuation function as in (1) with the same distribution for u_i. However, we can now make use of the variable threshold value as follows:

(7) $$y_i = 1 \text{ if } Y_i > t_i$$
$$= 0 \text{ otherwise,}$$

so that

(8) $$\Pr(y_i = 1) = \Pr(Y_i > t_i) = \Pr(u_i > t_i - x_i'\beta)$$
$$= \Pr(u_i/k > (t_i - x_i'\beta)/k)$$
$$= 1 - \Pr(\eta_i < (t_i - x_i'\beta)/k).$$

With this modification, the log likelihood function can now be written as:

(9) $$\log L = \Sigma - y_i \log\{1 + \exp[(t_i - x_i'\beta)/k]\}$$
$$+ (1 - y_i) \log\{\exp[(t_i - x_i'\beta)/k]/(1 + \exp[(t_i - x_i'\beta)/k])\}.$$

As before, this can be simplified to

(10) $$\log L = \Sigma (1 - y_i)[(t_i - x_i'\beta)/k] - \log\{1 + \exp[(t_i - x_i'\beta)/k]\}.$$

The presence of t_i allows k to be identified, which then allows us to isolate β so that the underlying fitted valuation function can be determined. (Note that if $t_i = 0$, we again have the conventional logit likelihood function.)

The likelihood function in (11) can be optimized directly using a general nonlinear function optimization computer program.[5] This procedure will yield separate estimates of β and k (and their individual asymptotic standard errors). However, estimates of $-1/k$ and β/k can be obtained more conveniently from conventional packaged logit algorithms. If we simply include the threshold, t_i, among the "explanatory" variables in an ordinary (maximum likelihood) logit model (as has typically been done by earlier researchers using CECV data), it is easy to see that:

$$(11) \qquad -\ (t,x') \begin{bmatrix} -1/k \\ \beta/k \end{bmatrix} = -x^{*\prime}\gamma^{*},$$

The augmented vectors of variables, x^* and coefficients, γ^*, may be treated as one would treat the explanatory variables and coefficients in an ordinary logit estimation. From γ^*, it is possible to compute point estimates of the desired parameters β and k. If we distinguish the elements of γ^* as $(\alpha, \gamma) = (-1/k, \beta/k)$, then $k = -1/\alpha$ and $\beta_j = -\gamma_j/\alpha$. However, accurate asymptotic standard errors for these functions of the estimated parameters cannot be computed directly. One task of this paper is to examine some alternative, relatively simple, methods for deriving these standard errors, using only the information gleaned from a conventional logit model.

One alternative is to use Taylor series approximation formulas for the variances of the desired parameters (Kmenta [8, p. 444]):

$$(12) \qquad \begin{aligned} \mathrm{Var}(k) &= \mathrm{Var}(-1/\alpha) = [1/\alpha^2]^2\ \mathrm{Var}(\alpha) \\ \mathrm{Var}(\beta_j) &= [\gamma_j/\alpha^2]^2\ \mathrm{Var}(\alpha) + [-1/\alpha]^2\ \mathrm{Var}(\gamma_j) \\ &\quad + 2\ [\gamma_j/\alpha^2][-1/\alpha]\ \mathrm{Cov}(\alpha,\gamma_j) \end{aligned}$$

A second possibility is to use the analytical formulas for the Hessian matrix corresponding to the likelihood function in (10) with the optimal values of β and k derived from γ^*. The negative of the inverse of this matrix can be used to approximate the Cramer-Rao lower bound for the variance-covariance matrix for β and k. Alternately, the *expected values* of the Hessian matrix elements are sometimes used in this process.[6] Formulas for the elements of the Hessian matrix and their expectations are provided in the Appendix.

In contrast, if the estimates of β and k are obtained directly by maximizing the log-likelihood function in (11), packaged programs for optimizing nonlinear functions will usually offer as an option the computation of the negative of the inverse of the Hessian matrix followed by a printout of the computed asymptotic standard errors or t-test statistics. Whichever way the asymptotic standard errors are determined, they are necessary for hypothesis testing regarding the signs and sizes of individual β_j parameters, an important objective of the modeling process.[7]

III. ILLUSTRATION: VALUING A RECREATIONAL FISHING DAY

We will illustrate our procedures with a highly simplified version of the valuation model (and a subset of the data) used in our earlier papers based on CECV data assuming normal errors. Our abbreviated sample consists of 1033 responses to an in-person survey of recreational salmon fishermen returning from fishing excursions.[8] After assessing the actual level of incidental expenses for the fishing day, respondents were asked if they would still have gone fishing if the fishing day had cost some (randomly assigned) number of dollars more. As explanatory variables in this simple illustration, we will use NFISH (the number of salmon caught), LGFISH (the weight of the largest fish, in pounds), TEMP (mean temperature that day, in degrees Celsius), PRECIP (total precipitation, in millimeters), and a dummy variable, NONRES, (which takes on a value of one if the respondent is not a local resident, and is zero otherwise). Table I summarizes the data.

Maximum likelihood estimates of β and k (and their asymptotic t-test statistics) from a very simple log-linear[9] version of the model in equation (10) are given in Table II. Also shown, for comparison, are point estimates of β and k computed from estimates of γ^* from an ordinary logit model (and approximate asymptotic t-ratios using standard errors computed by the Taylor series approximation given in (12) and from the expected value of the Hessian matrix).

In the past, empirical work using binary choice logit models with CECV data overlooked this reparameterization and therefore faced the awkward limitation of working only with *probability* estimates. This restricted the generality of the *WTP* functions unnecessarily. Here, however, any specification which could be estimated by OLS if *Y* were known can now be estimated quite cheaply with only CECV responses, and the estimated coefficients (other than k) can be interpreted *roughly* as one would interpret the results from an OLS regression..

The βs in this very simple specification give the percent change in *WTP* for a one-unit change in the level of each explanatory variable. As expected, the number and size of fish caught both increase the respondent's *WTP*: an additional fish increases *WTP* by about 9.16% (the average effect of an additional fish is $1.85). If the largest fish is one pound heavier, *WTP* is higher by about 2.27% (or by about $0.46 on average). The inverse of TEMP might be considered as a proxy for the avidity of the fishermen--only serious fishermen will be out on bad days. Therefore, it is not surprising that higher values of TEMP imply a less valuable fishing day. When the temperature is higher by 1 degree Celsius, the *WTP* is lower by about 3.2% (or by $0.65 on average). PRECIP will not affect the fish, but will definitely make the fishermen less comfortable: for every millimeter of rain, *WTP* is decreased by about 5.91% (or by about $1.19 on average). Non-residents have travelled further to go fishing, and can hence be expected to value the experience more (by about 58%, or $11.71 on average).

Clearly, the distortion in the standard error estimates due to use of the Taylor series approximation is very small in this example. This suggests that the model can quite readily be estimated using existing computer programs for conventional logit models. General nonlinear optimization programs are less common and typically require greater programming skill, so this result is very reassuring. We also provide estimates of the asymptotic t-ratios based on standard errors computed using the expectation of the Hessian matrix. These t-values appear to be biased downwards in this application.

The within-sample goodness-of-fit of the estimated model will often be of interest. As in our earlier studies, we can adapt some measures traditionally used with discrete choice models: (a.) individual prediction success, and (b.) aggregate prediction success.[10] These are displayed in Table III. Model "validation" would be analogous to the validation of OLS regression models.[11]

Compared to conventional methods for estimating models using closed-ended contingent valuation data, the ability to determine (systematically and simply) the effect upon expected *WTP* of changes in the levels of each explanatory variable is the clear advantage of this new approach. Nevertheless, we might still be interested in the marginal mean of the distribution of individual *WTP* values. Here, we can make use of an identity which is familiar to regression analysts. Since the mean of u_i is zero at the optimal parameter values, $E(Y_i) = E(x_i'\beta + u_i) = E(x_i'\beta) + E(u_i) = E(x_i'\beta)$.[12] Mean fitted *WTP* across all respondents is \$20.20.

It is useful to compare the results obtained for this logistic model (Tables II and III) with those which *would* have been obtained if normal errors had been assumed instead. (See our earlier paper, Cameron and James [5]) for the derivation of this procedure, which is analogous to that presented in section II except that the different densities results in different likelihood functions.) In both cases, the underlying fitted *WTP* is given by $x_i'\beta$. If the same functional form is imposed for the relationship between the unobservable Y_i and $x_i'\beta$, the parameters β will be comparable. Furthermore, the standard deviation of the fitted error distribution in the logistic case will be $b = k\pi/\sqrt{3}$, which will be comparable to σ in the normal case.

Table IV therefore shows results for the same log-linear specification, but in this case we assume normal errors.[13] Note first of all that the maximized value of the log-likelihood for the model with logistic errors is considerably better than that for the normal errors (-388.4 versus -403.9). However, the models are not nested, so we cannot assess whether this difference is statistically significant. Still, the superiority of the logistic model is corroborated by the fact that the error standard deviation for the normal model is 0.7591, whereas for the logistic model, it is only 0.6982. Within-sample goodness-of-fit for the model with normal errors appear in Table V. These measures also suggest that in this case, the logistic model is probably preferred.

The point estimates from the two models do differ slightly. The implications for policy of the different error assumptions are best seen by examining the means of the implied derivatives of *WTP* with respect to each variable--these are displayed in Table VI. Also, while the overall *marginal* mean of the fitted values for *WTP* with logistic errors was $20.20, it is $19.69 for the normal model.

IV. RECONSIDERING THE ESTIMATES OBTAINED IN EARLIER LOGIT ANALYSES

Bishop and Heberlein [1] and Bishop *et al.* [2] addressed the valuation of goose hunting permits for the Horicon Zone (in east central Wisconsin). They employ a simple logit model to analyze respondents' willingness-to-accept (*WTA*) compensation for their permit. In their initial model, with only the logarithm of the offered amount as an explanatory variable, the log-odds (LO) of the probability of accepting an offer to sell (namely, the fitted value of $x_i^{*\prime}\gamma^*$) is

(13) $LO_i = 3.24 - .74 \log(DOLLARS_i)$.

Solving for the β parameters yields the underlying valuation function:

(14) $\log(WTA_i) = 4.378$.

The fact that the fitted value of the logarithm of *WTA* is just a constant means that we can recover a point estimate for individual value of $79.68. Whereas Bishop and Heberlein report that the expected value of a permit calculated from this model is $101 (versus $63 for their actual cash offers in simulated markets), our interpretation of their fitted model yields a mean value that is much closer to that found in actual transactions.

Bishop and Heberlein's second model includes a categorical "commitment" variable--a four-item attitude scale expressing the level of commitment each hunter had to goose hunting with larger values expressing greater commitment. Bishop and Heberlein assume that *WTA* is linear in the arbitrarily assigned levels of this categorical variable (a set of three dummy variables might have been preferable). Their estimated model is:

(15) $LO_i = -.58 - .84 \log(DOLLARS_i) + .40\ COMMITMENT_i$.

After transformation, the fitted *WTA* relationship becomes:

(16) $\log(WTA_i) = -.6905 + .4762\ COMMITMENT_i$.

If we knew the mean "level" of the COMMITMENT variable, we could substitute this number into the formula in (16) and compute the value of log(*WTA*) at the "mean" of the data. This number would correspond to the overall marginal mean of *WTA* in the sample, based on the fitted model. Without the data on COMMITMENT, however, we cannot reconstruct a value to compare to Bishop and Heberlein's number.

In the Sellar, Stoll, and Chavas [12] study, respondents were asked to respond "yes" or "no" to the following question: "If the annual boat ramp permit cost $X in 1980, would you have purchased the permit so that you could have continued to use the lake throughout the year?" These authors used a simple logit model to estimate the probability that the respondent would answer "no" to a given value of X. In a second paper employing a subset of the same dataset, [13], these same authors devote considerable attention to applying the tradition of random utility maximization models developed by McFadden [10]. With the methodology described in the present paper, however, it becomes apparent that appeals to this tradition are unnecessary. The data environment here is distinctly different from McFadden's discrete choice framework. CECV data are much richer. The "valuation functions" fitted to CECV data are simply inverse demand functions. There is a mature literature on the types of demand functions which satisfy the regularity conditions necessary to render them consistent with classical utility maximization. These functions can be adopted to whatever extent the range of available variables will allow.

In the process of untangling the underlying valuation functions employed by Sellar, Stoll, and Chavas [12] it is more convenient for us to work *not* with the probability of a "no" response, but instead with the fitted probability that the respondent would answer "yes" (implying that their valuation is larger than the threshold value)[14]. Additional data include the number of visits per year, q. Let LO be the log of the odds of responding "yes" to the *WTP* question. The models are fitted for four different locations:

(17) Conroe: $LO_i = 6.13 - 1.79 \log(X_i) + .16 \log(q_i)$

Livingston: $LO_i = 3.06 - 1.37 \log(X_i) + .67 \log(q_i)$

Somerville: $LO_i = 4.78 - 1.26 \log(X_i) + 1.75 \log(q_i)$

Houston: $LO_i = 2.32 - 0.99 \log(X_i) + .47 \log(q_i)$

Transforming these coefficients to uncover the β values yields an underlying fitted *WTP* relationship for each location:

(18) Conroe: $\log(WTP_i) = 3.42 + 0.089 \log(q_i)$

Livingston: $\log(WTP_i) = 2.23 + 0.489 \log(q_i)$

Somerville: $\log(WTP_i) = 3.79 + 1.389 \log(q_i)$

Houston: $\log(WTP_i) = 2.34 + 0.475 \log(q_i)$

Thus, log(*WTP*) is a function of log(q). If we interpret *WTP* as the product of the average price willingly paid (p) times the number of units (q), these equations all have the form:

(19) $$\log(pq) = \log(p) + \log(q) = a + b \log(q).$$

We can rearrange these formulas to isolate log(q) on the left-hand side:

(20) $$\log(q) = [a/(1-b)] - [1/(1-b)] \log(p).$$

We then arrive at point estimates for the implied demand functions for each of the four areas:

(21) Conroe: $\log(q) = 3.76 - 1.10 \log(p)$

Livingston: $\log(q) = 4.37 - 1.96 \log(p)$

Somerville: $\log(q) = -9.76 + 2.57 \log(p)$

Houston: $\log(q) = 4.46 - 1.90 \log(p)$

The coefficients on log(p) can be interpreted as price elasticities of demand for boat ramp use. As noted by the authors, the results for Somerville are inconsistent with the theoretical notion that demand curves ought to slope downward, and therefore this model should probably be rejected. For the other areas, demand for boat ramp permits would appear to be uniformly somewhat elastic, but to vary across locations. If we had the parameter covariance matrices for the original models in (17), it would be a simple matter to compute approximate standard errors for these elasticity point estimates using the Taylor series approximation described in section II.

V. CONCLUSIONS

Extensive evaluation of CV methods in experimental settings suggests that they are quite reliable, despite earlier concerns about the potential for "strategic bias" and other hazards. These survey instruments are becoming increasingly popular. Mitchell and Carson mention approximately 125 citations in their partial listing of contingent valuation studies.

Our earlier research described CECV estimation procedures which conformed as closely as possible with OLS regression models, but which could be expensive or time-consuming to compute, since they required numerical evaluation of integrals which were not of closed form. The statistical model described here adapts our earlier work to the assumption of an underlying logistic distribution for the conditional density of the unobserved valuation. Since computations in this case involve only ratios of exponentiated terms, computation can be significantly cheaper, especially for large samples with large numbers of explanatory variables and for complex functional relationships between valuation and these variables.

This logistic error model for CECV data can be applied quite simply by anyone who has access to a conventional maximum likelihood binary logit computer algorithm. Selecting from a range of alternatives, we have demonstrated two techniques for arriving at approximate asymptotic standard errors: by Taylor's series expansion and using transformed ordinary logit point estimates in the formulas for the expected value of the Hessian matrix corresponding to the full likelihood function for our logistic model. The first, simpler, method seems to be sufficient (at least in our illustration). Compared to the model with normal errors, the logistic model can be expected to generate slightly different point estimates (and therefore somewhat different policy implications), but the divergence does not seem to be too severe. After all, given the similarities in the shapes of the simple distributions, one would expect comparable results.

In our abbreviated empirical application, we find that for the simple log-linear specification and our limited set of explanatory variables, the logistic error model achieves a higher maximized log-likelihood value and results in a smaller fitted error standard deviation than does the normal error model. However, this will not always be the case--it will depend on the data.

This logistic procedure also allows us to go back and reinterpret some earlier results generated by other researchers, since the derivation of this model brings out the correct interpretation of the CECV parameter estimates yielded by simple logit discrete choice models. It is easy to recover the underlying demand functions with no more than just the fitted models reported in the published versions of these papers. If the researchers who performed these earlier studies have retained the covariance matrices for their parameter estimates, it would be possible for them to calculate approximate asymptotic standard error estimates without repeating any of the original computations.

A tangential result is that it is probably unnecessary for researchers using CECV data to appeal to McFadden's random utility maximization models for discrete choice situations. Conventional demand theory would seem to suffice.

REFERENCES

1. R.C. Bishop and T.A. Heberlein, Measuring values of extramarket goods: are indirect measures biased? *Amer. J. Agr. Econom.* **61**, 926-930 (1979).

2. R.C. Bishop, T.A. Heberlein, and M. J. Kealy, Contingent valuation of environmental assets: comparisons with a simulated market, *Natural Resources J.* **23**, 619-633 (1983).

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

4. T.A. Cameron and M.D. James, Efficient estimation methods for use with 'closed-ended' contingent valuation survey data," forthcoming, *Rev. Econom. Statist.* (1987a).

5. T.A. Cameron and M.D. James, "Estimating Willingness-to-Pay from Survey Data: An Alternative Pre-test Market Evaluation Procedure," revision of Department of Economics Discussion Paper #415, University of California, Los Angeles (1987b).

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

7. N.A.J. Hastings and J.B. Peacock, "Statistical Distributions," Wiley, New York (1975).

8. J. Kmenta, "Elements of Econometrics," Macmillan, New York (1971).

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

10. D. McFadden, Quantal choice analysis: A survey, *Ann. Econom. Soc. Measure.* **5**, 363-390 (1976).

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

12. C. Sellar, J.R. Stoll and J.P. Chavas, Validation of empirical measures of welfare change: a comparison of nonmarket techniques, *Land Econom.* **61**, 156-175 (1985).

13. C. Sellar, J.P. Chavas, and J.R. Stoll, Specification of the logit model: the case of valuation of nonmarket goods, *J. Environ. Econom. Management* **13**, 382-390 (1986).

TECHNICAL APPENDIX

Using the notation established in the text, we first define the following simplifying abbreviations:

$$\eta_i = (t_i - x_i'\beta)/\sigma \qquad R_i = 1/(1+\exp(-\eta_i)) \qquad S_i = R_i^2\exp(-\eta_i)$$

The gradient vector for this model is then given by:

$$\partial log\ L/\partial\beta_r = \Sigma\ (x_{ij}/k)\ \{(y_i - 1) + R_i\ \} \qquad r = 1,\ldots,p$$

$$\partial log\ L/\partial\sigma = \Sigma\ (\eta_i/k)\ \{\ (y_i - 1) + R_i\ \}$$

The elements of the Hessian matrix are:

$$\partial^2 logL/\partial\beta_r\partial\beta_s = -(1/k^2)\ \Sigma\ x_{ij}x_{ik}\ S_i \qquad r,s = 1,\ldots,p$$

$$\partial^2 logL/\partial\beta_r k = -(1/k)^2\ \Sigma\ x_{ij}\ \{\ (y_i - 1) + R_i(1 + \eta_i)\ \} \qquad r = 1,\ldots,p$$

$$\partial^2 logL/\partial k^2 = -(1/k^2)\ \Sigma\ (2\eta_i)\ \{\ (y_i - 1) + R_i\ \} + \eta_i^2 S_i$$

There exist function optimization algorithms which will find the optimal parameter values using only the function itself (and numeric derivatives). However, analytic first (and second) derivatives can sometimes reduce computational costs considerably.

The expectation of y_i is $[1/(1+\exp(\eta_i))]$. The negatives of the expectations of the Hessian elements are as follows:

$$-\ E(\partial^2 logL/\partial\beta_r\partial\beta_s) = (1/k^2)\ \Sigma\ x_{ij}x_{ik}\ S_i \qquad r,s = 1,\ldots,p$$

$$-\ E(\partial^2 logL/\partial\beta_r k) = (1/k^2)\ \Sigma\ x_{ij}\eta_i\ S_i \qquad r = 1,\ldots,p$$

$$-\ E(\partial^2 logL/\partial k^2) = (1/k^2)\ \Sigma\ \eta_i^2\ S_i$$

Table I

Descriptive Statistics (n = 1033)

Variable Name	Description	Mean (proportion)	Standard Deviation
t_i	offered threshold	27.78	23.77
y_i	"yes" (willing to pay t_i)	0.4869	
NFISH	number of fish caught	0.9671	1.476
LGFISH	weight of largest fish	3.191	5.404
TEMP	mean temperature (C)	13.11	4.096
PRECIP	total precipitation (mm)	0.7780	2.340
NONRES	1 if non-resident	0.06583	

Table II

Estimation Results: Logistic Errors

"Dependent" Variable = Log(Unobserved Willingness-to-Pay)

Variable	Maximum Likelihood Parameter Estimates[a]		Parameters Computed from Logit Model Estimates		
int	3.215	(24.49)	3.215[b]	(26.49)[c]	(20.52)[d]
NFISH	0.09163	(3.100)	0.09163	(3.100)	(2.402)
LGFISH	0.02267	(3.342)	0.02267	(3.342)	(2.589)
TEMP	-0.03198	(-3.815)	-0.03198	(-3.815)	(-2.955)
PRECIP	-0.05913	(-3.126)	-0.05913	(-3.126)	(-2.421)
NONRES	0.5796	(4.292)	0.5796	(4.292)	(3.325)
k	0.3850[e]	(15.65)	0.3850	(15.65)	(12.12)

Maximized log-likelihood: -388.4

[a] Asymptotic t-ratios from maximum likelihood estimation in parentheses.

[b] Due to the invariance property of maximum likelihood estimators, the point estimates should be identical by either method.

[c] Approximate t-ratios constructed from Taylor series standard error approximations.

[d] t-ratios using standard errors computed from inverse of negative of expected value of Hessian matrix corresponding to the log-likelihood function in (8).

[e] standard deviation of error distribution is therefore 0.6982.

Table III

Prediction Success: Logistic Errors

Individual:

		Actual	
		Would Pay t_i	Would Not
Predicted	Would Pay t_i	415	76
	Would Not	88	454

(Percent Correct: 84.12)

Aggregate:

	Predicted	Actual
Would Pay t_i	503.0	503
Would Not	530.0	530

Table IV

Estimation Results: Normal Errors

"Dependent" Variable = Log(Unobserved Willingness-to-Pay)

Variable	Maximum Likelihood Parameter Estimates[a]		Parameters Computed from Probit Model Estimates		
int	3.210	(23.95)	3.210[b]	(23.95)[c]	(30.31)[d]
NFISH	0.06976	(2.392)	0.06981	(2.393)	(2.759)
LGFISH	0.02760	(3.983)	0.02759	(3.981)	(4.453)
TEMP	-0.03342	(-3.586)	-0.03342	(-3.586)	(-4.552)
PRECIP	-0.06113	(-3.029)	-0.06112	(-3.030)	(-3.882)
NONRES	0.6304	(4.210)	0.6302	(4.210)	(5.296)
σ	0.7591	(18.00)	0.7589	(17.96)	(21.32)
Maximized log-likelihood:	-403.9				

[a,b,c] See Table II.

[d] t-ratios using standard errors computed from inverse of negative of expected value of Hessian matrix corresponding to the log-likelihood function for the model with normal errors. (See Cameron and James [5].)

Table V

Prediction Success: Normal Errors

Individual:		Actual	
		Would Pay t_i	Would Not
Predicted	Would Pay t_i	407	72
	Would Not	96	458

(Percent Correct: 83.74)

Aggregate:		Predicted	Actual
	Would Pay t_i	498.7	503
	Would Not	534.3	530

Table VI

Comparison: Logistic Model versus Normal Model

Mean Derivatives[a] (Standard Deviations) of *WTP*

Variable	Logistic Model Estimates		Normal Model Estimates	
NFISH	$ 1.851	(0.7105)	$ 1.374	(0.5454)
LGFISH	0.4580	(0.1758)	0.5436	(0.2158)
TEMP	-0.6460	(0.2480)	-0.6582	(0.2613)
PRECIP	-1.194	(0.4585)	-1.204	(0.4780)
NONRES	11.71	(4.494)	12.41	(4.929)

[a] Since $\partial WTP/\partial x_j = \beta_j * \exp(x_i'\beta)$, these derivatives vary across i. We compute the fitted derivative for each observation, and report the means and standard deviations of these quantities across observations.

[1] The data used in this study were collected under a project sponsored by the Department of Fisheries and Oceans, Vancouver, Canada, supervised by Michelle D. James, my co-author on earlier related papers. The comments of four anonymous referees on these earlier papers have also contributed to the evolution of the present paper. E. Jane Murdoch has offered helpful suggestions.

[2] Or, over the curve, depending upon whether one is focusing on the "yes" or "no" response.

[3] These models can be adapted very simply to accommodate willingness-to-accept (*WTA*).

[3] Note that many textbooks (e.g. Maddala [9]) exploit the symmetry around zero of the standard logistic distribution to simplify these formulas even further. We simplify this way to preserve consistency with the next model where we estimate k explicitly.

[5] We used a program called GQOPT.

[6] The outer product of the gradient vector evaluated at the optimum is sometimes used. However, since the expectation of the Hessian has simple formulas, it is probably preferred in this application.

[7] Of course, if estimates are achieved by optimization of (10), hypothesis testing regarding the βs is the same as in any maximum likelihood context.

[8] This is the Victoria, British Columbia, subsample.

[9] If our underlying valuation function is assumed to have $\log(Y_i)$ on the left side, then we simply substitute $\log(t_i)$ wherever t_i appears in the formulas.

[10] For individual prediction successes, one counts up the number of times the model produces a fitted probability greater than (less than) 0.5 that the respondent will pay the stipulated price when the individual actually responds that he *would* (would not) pay that amount. For "aggregate prediction success," each respondent is assumed to represent some large equal number of respondents with identical characteristics and fitted choice probabilities are viewed as proportions. Fitted probabilities are summed and compared to actual responses.

[11] For example, out-of-sample prediction success can be assessed by fitting the model to a randomly-chosen subset of the data and observing its predictive capability for the rest of the dataset.

[12] In OLS, this property is borne out by the fact that the fitted regression line always passes through the means of the data.

[13] Readers familiar with our other work using excerpts from the same dataset (in particular, Cameron and James [5]) will note that fitted mean values of *WTP* can also be quite sensitive to the choice of linear versus log-linear functional forms. A Box-Cox generalization which subsumes these two alternatives is examined in Cameron and James [4] for the full sample.

[14] The probability of a "no" answer is $1/[1+\exp(-x^{*\prime}\gamma^*)]$; the probability of a "yes" answer is just $1/[1+\exp(x^{*\prime}\gamma^*)]$, so the change merely alters the sign on all coefficients γ^*.