ESTIMATION USING CONTINGENT VALUATION DATA FROM A "DICHOTOMOUS CHOICE WITH FOLLOW-UP" QUESTIONNAIRE

by

Trudy Ann Cameron*
Department of Economics
University of California
Los Angeles, CA 90024-1477
(310)825-3925 FAX (310)825-9528
Internet: CAMERON@ECON.SSCNET.UCLA.EDU

and

John Quiggin
Department of Economics
Research School of Social Sciences
Australian National University
Canberra, Australia

Cameron, T. A. and Quiggin, J. (1994), ‘Estimation using contingent valuation data from a "Dichotomous choice with follow-up" questionnaire’, *Journal of Environmental Economics and Management* 27(3), 218–34.

* We would like to acknowledge the helpful comments of two anonymous reviewers and of participants at the Southern California Environmental Economics Workshop held at UC Santa Barbara, May, 1993.

C:\WP51\DOC\DCF.DOC

ESTIMATION USING CONTINGENT VALUATION DATA FROM A "DICHOTOMOUS CHOICE WITH FOLLOW-UP" QUESTIONNAIRE

ABSTRACT

Dichotomous choice (referendum) contingent valuation questions have gained popularity over the last several years due to their purported advantages for avoiding many of the biases known to be inherent in other value elicitation formats. However, this type of valuation question is inefficient in that a much larger number of observations is required to identify the distribution of values with any degree of accuracy. An alternative questioning strategy introduces a second value threshold which elicits a second discrete response. Previous analyses of these double-bounded referendum surveys have maintained the hypothesis that the *identical* unobserved resource value motivates both responses. We relax this assumption and find that while the implied resource values deduced from the first and second responses may be drawn from the same distribution, and the implied values are highly correlated, the values elicited by each question are definitely not identical. Furthermore, assuming that they are identical severely distorts the estimated valuation distribution.

C:\WP51\DOC\DCF.DOC

ESTIMATION USING CONTINGENT VALUATION DATA FROM A "DICHOTOMOUS CHOICE WITH FOLLOW-UP" QUESTIONNAIRE

1. Introduction

Dichotomous choice contingent valuation questions have gained popularity over the last several years. This is due primarily to their purported advantages in avoiding many of the biases known to be inherent in other formats used in the contingent valuation (CV) method. Two standard references which discuss different CV techniques are Cummings, Brookshire, and Schulze (1986) and Mitchell and Carson (1989). Whereas several varieties of bias may be minimized by dichotomous choice valuation questions, this elicitation method can be highly statistically inefficient in that vastly larger numbers of observations are required to identify the underlying distribution of resource values with any given degree of accuracy.

An alternative questioning strategy, intended to reduce this inefficiency, was first proposed and implemented by Carson, Hanemann, and Mitchell (1986). They advocate introducing a second offered threshold in a "follow-up" dichotomous choice CV question which elicits a second discrete response. In practice, if a respondent indicates a willingness to pay the first offered amount, the new threshold is about double the first one. If the respondent is unwilling to pay the first offered amount, the second threshold is reduced to about half the original amount. This questioning strategy has also been called a "double-bounded referendum" approach.

Arrow et al. (1992) advocate discrete choice contingent valuation questions over other formats in their assessment of the reliability of CV techniques for quantifying passive use values in the context of oil spills. However, they note parenthetically that "If a double-bounded dichotomous choice or some other question form is used in order to obtain more information per respondent, experiments should be developed to investigate biases that may be introduced." (p. 52). This research addresses these possible biases for this double-bounded case.

Carson and Mitchell (1987) employ survival analysis statistical techniques to analyze dichotomous choice with follow-up data. These methods were originally conceived to handle product failure data collected at irregular intervals. Much of this literature has emphasized Weibull distributions for the variable in question. Hanemann, Loomis, and Kanninen (1991) use maximum-likelihood models to analyze double-bounded referendum contingent valuation data under an assumption of normality. Both of these papers, however, maintain the hypothesis that a single implicit true valuation drives respondents' answers to both of the questions in this survey format. This paper proposes a more-general maintained hypothesis. Our estimation method allows the valuation information elicited at each of the two stages to be the same, or different, as the data dictate.

In analyzing data from a dichotomous-choice-with-follow-up (DCF) questionnaire, it is certainly important for the researcher to acknowledge explicitly the endogeneity of the second offered amount. Using a sample of data from an actual DCF survey, we examine the distortions in the final value estimates which can be introduced when restrictive conventional assumptions are imposed. In our new specifications, separate distributional parameters for willingness to pay (for the two CV questions), as well as the correlation across the two questions in the two true underlying unobserved values, are estimated explicitly. Our models allow the researcher to test statistically for the equivalence of the implied valuation distributions across the original and follow-up questions. They also allow rigorous tests of restrictions that might be imposed upon both the distribution parameters and the error correlation.

We determine that the usual assumption--that identical value distributions are elicited by the first and follow-up questions--is implausible for our data set. This assumption also appears to compound the problem of the implicit distribution of value estimates being influenced by the starting point (i.e. the first bid). The usual assumption also precludes an assessment of first response effects.

This paper is organized as follows. Section 2 outlines our new model based on latent bivariate normal errors for a two-equation system of discrete variables. Section 3 describes a convenience sample of data to demonstrate this model. Section 4 gives empirical results for our model and its special cases. Section 5 reviews the implementation of the standard assumptions and their consequences when employed with these data. Section 6 explores our data for evidence of "first response" effects on values implied in the second round of questioning and section 7 looks at the implications of the two classes of models for detection of "starting point" effects in the data. We briefly cover the quantitative implications of our models for the value of the specific resource in question in section 8, and section 9 concludes.

2. A General Model for "Dichotomous Choice with Follow-up" Survey Data

Throughout this analysis, we will be emphasizing the importance of plausible stochastic assumptions in the estimation of valuation models. If our objective in this paper was to ascertain the best possible estimate of the value of a specific environmental resource, we would employ covariates and go on to explore a vastly wider array of functional forms for the systematic portions of our valuation functions. The example employed here should be viewed simply as illustrative. Nevertheless, it succinctly conveys the importance of the issue we are highlighting.

As a point of departure for the model to be developed in this paper for DCF data, we rely upon the econometric framework developed in Cameron and James (1987) and in Cameron (1988). Those papers show how the data collected using a single discrete-choice CV question can be employed in censored dependent variable models having a natural regression-like interpretation. These models can be estimated directly using general maximum likelihood optimization algorithms, or, identical results can be calculated from the output of packaged probit algorithms.

Much of the extant empirical work using discrete-choice CV data *without* a follow-up has assumed, for expedience, that the underlying error distribution is logistic. This assumption leads to convenient closed-form integrals for the cumulative probability density functions which must be evaluated. For the present problem, however, we prefer normality of the errors since we wish to model each participant's two discrete responses jointly. Bivariate normal probability density functions are the most familiar bivariate distributions employed commonly by statisticians and their properties are well-understood. Crucially, they allow for a non-zero correlation, whereas the standard logistic distribution does not.

Hanemann (1991) does suggest a parametric test of the consistency between the responses to the second and first bids in his comments upon the work described in Imber, Stevenson and Wilks (1991). He recommends conducting

> "...a parametric, likelihood ratio test of the overall consistency of the responses to the first and second bids by estimating two models, one based exclusively on the responses to the first bid (e.g. a conventional logit model...) and the other based exclusively on the responses to the second bid (this involves maximizing the likelihood function framed in terms of the conditional probabilities for the response to the second bid), and then comparing the sum of the log-likelihoods with the log-likelihood function obtained from estimating the combined first and second bids (i.e. the model...actually estimated...)."

Since Hanemann's proposed method utilizes the stochastic assumptions of the usual logit specification, however, the error correlations in this model would not be estimable. Thus his approach would not be equivalent to that proposed in this paper.

For our model with normally distributed errors, assume that each respondent has some unobserved true point valuation for the environmental resource in question at the moment the first dichotomous choice CV question is posed. Let this unobserved value be y_{1i}. Let the first offered threshold, assigned arbitrarily to this individual, be denoted by t_{1i}. We will assume that the individual will state that they are willing to pay the offered amount (I_{1i} = 1) if $y_{1i} \geq t_{1i}$. They will be unwilling to pay this much ($I_{1i} = 0$) if $y_{1i} < t_{1i}$. Now let the unobserved valuation y_{1i} consist of a systematic component, $x_{1i}'\beta_1$, which is a function of a vector x_{1i} of observable attributes of the respondent, x_{1i}, plus an unobservable random component, ε_{1i} (distributed $N(0,\sigma)$), which absorbs all unmeasured determinants of the value of the resource to this individual. The discrete response indicator variable I_{1i} is the single endogenous (dependent) variable in this framework. Analysis of the first-stage responses is facilitated by the fact that the offered values in the first round, t_{1i}, are assigned randomly and therefore have no possible have no possible correlation, ex ante, with the error terms, ε_{1i}.

So far, the development of the model mimics exactly the specifications used for single-threshold CV data. But in the typical DCF framework, once an individual has been randomly assigned their initial offered value, the follow-up offer will take on one of two alternative predetermined values (one higher and one lower) depending upon the response to the first question. The probability of receiving the predetermined *higher* offer is just the probability of responding "yes" to the first WTP question. The probability of receiving the predetermined *lower* offer is the same as the probability of a "no" response to the first WTP question. The second offered threshold is clearly *not* independent of valuation information which the respondent has revealed in answering the first WTP question.

In particular, it would be highly inappropriate to use just the set of second-round responses from a DCF questionnaire in a model which employed the same assumptions as in the Cameron and James (1987) or Cameron (1988) papers. It would also be invalid to simply "pool" all of the thresholds and responses from the first and second questions in the estimation of a single valuation function. The endogeneity of the second offered amount precludes either of these approaches.

Adopting similar notation for the data pertaining to the follow-up question, let y_{2i} be the respondent's implicit underlying point valuation of the resource at the moment the follow-up question is posed. Crucially, this may or may not be identical to y_{1i}, depending upon the presence or absence of a strategic adjustment or upon other effects of the questioning strategy experienced by the respondent. The indicator variable, I_{2i}, will be one if $y_{2i} \geq t_{2i}$ (where t_{2i} is explicitly endogenous), and zero if $y_{2i} < t_{2i}$. The point valuation will again be assumed to consist of a systematic component, $x_{2i}'\beta_2$ (where in general x_2 need not be identical to x_1), and a random unobservable component, ε_{2i}, distributed $N(0,\sigma)$. Failure to acknowledge that ε_{2i} is correlated with ε_{1i} (and thus also with t_{2i}) will result in potentially serious endogeneity bias in all of the coefficients comprising the vector β_2.

A viable strategy for dealing with this endogeneity is to view the estimation problem as analogous to the more-common problem posed by a system of two equations with correlated errors. In linear systems, econometricians are familiar with recursive systems of two equations where the second equation contains the first equation's dependent variable on its right-hand side. In that context, if the errors can be assumed to be independent, each equation can be estimated separately by ordinary least squares. If this assumption is not tenable, systems estimation methods must be employed.

In the present situation, however, we are not dealing with the familiar case of continuous dependent variables, but the two discrete responses, I_{1i} and I_{2i}. The offered threshold entering into the second "equation" reflects the probabilities associated with the discrete outcome I_{1i}. We must therefore develop the model in the context of the joint distribution of (y_{1i},y_{2i}). We will assume a bivariate normal distribution, $BVN(x_1'\beta_1,x_2'\beta_2,\sigma,\sigma,\rho)$ for these two implicit valuations. There are four possible pairs of responses to these questions: (I_{1i},I_{2i}) = (1,1), (1,0), (0,0) and (0,1). Dropping the i subscripts for ease of exposition, recall that $I_1 = 1$ implies $y_1 \geq t_1$. Using $y_1 = x_1'\beta_1+\varepsilon_1$, this condition can be expressed equivalently as $(\varepsilon_1/\sigma_1) > (t_1-x_1'\beta_1)/\sigma_1$, where ε_1/σ_1 is a standard normal random variable. The analogous transformation can be applied to y_2 and t_2 in determining the formula for the probability function for outcome I_2.

Denote the standardized normal error ε_1/σ_1 as z_1 and denote ε_2/σ_2 as z_2. The analysis can proceed in terms of the probabilities associated with regions in the domain of a standard bivariate normal distribution where the pair (z_1,z_2) is distributed $BVN(0,0,1,1,\rho)$. To simplify the notation in preparation for writing the log-likelihood function for this model, let $g(z_1,z_2)$ be the bivariate standard normal density function. This density takes the explicit form:

(1) $$g(z_1,z_2) = [1/(2\pi(1-\rho^2)-)] \exp \{ -(2-2\rho^2)^{-1} [z - 2\rho z_1 z_2 + z] \},$$

where $z_1 = (t_1 - x_1'\beta_1)/\sigma_1$ and $z_2 = (t_2 - x_2'\beta_2)/\sigma_2$.

The log-likelihood function for the model then takes the following form.

(2) $$\text{LogL} = \Sigma_i \{ (I_1 I_2)\log [\qquad g(z_1,z_2)\ dz_2\ dz_1]$$
$$(t_1-x_1'\beta_1)/\sigma_1 \qquad (t_2-x_2'\beta_2)/\sigma_2$$

$$(t_1-x_1'\beta_1)/\sigma_1$$
$$+ (1-I_1)(I_2) \log [\qquad g(z_1,z_2)\ dz_2\ dz_1]$$
$$- \qquad (t_2-x_2'\beta_2)/\sigma_2$$

$$(t_1-x_1'\beta_1)/\sigma_1 \qquad (t_2-x_2'\beta_2)/\sigma_2$$
$$+ (1-I_1)(1-I_2) \log [\qquad g(z_1,z_2)\ dz_2\ dz_1]$$
$$- \qquad -$$

$$(t_2-x_2'\beta_2)/\sigma_2$$
$$+ (I_1)(1-I_2) \log [\qquad g(z_1,z_2)\ dz_2\ dz_1] \}.$$
$$(t_1-x_1'\beta_1)/\sigma_1 \qquad -$$

Note that most of the parameters to be estimated, β_1, β_2, σ_1, and σ_2, appear in the limits to the integrals. The remaining correlation parameter, ρ, is embedded in the $g(z_1,z_2)$ terms.

This general model can be readily estimated using standard packaged bivariate probit algorithms such as those offered in the LIMDEP computer program. Recall that models for *single*-threshold dichotomous choice can be estimated using conventional maximum likelihood probit algorithms. Exploiting the invariance property of maximum likelihood, the resulting standard probit parameter point estimates can then be transformed to yield an associated regression-like relationship for the dichotomous choice model. To obtain the variance-covariance matrix corresponding to the transformed parameters, one can take advantage of a formula offered in Lehmann (1983), pointed out by Patterson and Duffield (1991). Analogous transformations exist for the bivariate case.

3. A Brief Description of the Data

The literature concerning dichotomous choice contingent valuation methods is now quite substantial. One example is a study undertaken by the Australian Resource Assessment Commission (RAC) in 1990 (Imber, Wilks and Stevenson). The study was part of an evaluation of alternative proposals for the management of the Kakadu region of the Northern Territory (NT) in Australia. Kakadu is an important wilderness area, but also contains significant mineral deposits. The establishment of a proposed National Park was divided into three stages. At the time of the study, the first two stages had been declared as National Park, with the excision of an existing uranium mine. The debate concerned the proposed third stage, which included a promising site for gold and uranium mining at Coronation Hill.

The RAC was commissioned in 1990 to report to the Australian Government on the appropriate policy response. As part of this effort, Imber, Wilks and Stevenson (1990) undertook a contingent valuation of the preservation option. A total of 2034 adult Australians were interviewed. A leading polling agency was employed to conduct the interviews. The questionnaire was subject to extensive pre-testing, and careful attempts were made to cope with all of the sources of bias discussed in the literature.

Eight different "treatments" were used in the CV questionnaire, arising from four payment levels and two different scenarios (minor damage and major damage) describing possible damage to conservation values arising from mining. In each treatment, respondents were asked some introductory questions, then asked to nominate areas of major environmental concern. Only a small proportion (2 per cent) specifically named Kakadu at this stage.

Respondents were then asked about their knowledge of Kakadu, and were presented with photographs and maps describing the conservation zone. Each respondent was then presented with one of the two scenarios and asked two willingness-to-pay questions. If the first question was answered affirmatively, the amount was increased (approximately doubled), otherwise it was decreased (approximately halved). The payment vehicle was an increase in taxes.

After answering the WTP questions, subjects were asked for reasons why they were (or were not) willing to pay the amounts in question. They were then presented with a number of questions eliciting attitudes to environmental issues and a range of questions on socio-economic variables (age, sex, education, income, national origin, occupation).

The issue in question was politically contentious and a large proportion of respondents had strongly held views on the subject. Supporters of mining were unlikely to state any positive willingness to pay (and might well have indicated negative willingness to pay if asked). The modal explanation among subjects who answered "No" was a statement of the form "Support mining/good for the country." A smaller proportion of subjects gave explanations the form "Too much money/not worth it to me," which would indicate a willingness to pay a positive amount less than the threshold asked. Similarly, committed supporters of preservation were likely to answer "Yes" to questions involving even very high thresholds.

The data used in the present paper are drawn from the published aggregate responses in Imber, Wilks and Stevenson (1991). Attention was confined to the 1013 respondents presented with the "minor damage" scenario, which was considered by Imber, Wilks and Stevenson to reflect majority scientific judgment of the likely impact of the mine. The responses are summarized in Table 1.

4. Empirical Results under the General Model

If the researcher is merely trying to quantify the location and scale of the current distribution of valuations in a particular sample, additional covariates may not be required. The $x_1'\beta_1$ and $x_2'\beta_2$ terms will be simple intercept terms represented by the scalars β_1 and β_2. Of course, if the model is intended to be used for forecasting, simulation, or benefits transfer, the use of any available regressors will probably be advisable. Regressors are also crucial if one is attempting to ascertain the marginal value of changes in amenity levels associated with particular resources.

To emphasize the possible distortions stemming from the usual assumptions employed with DCF data, it is sufficient to simplify the model until it involves only the means, the variances, and the correlation of the assumed bivariate normal WTP values elicited by the initial and follow-up dichotomous choice CV questions on our survey. (All of the procedures we utilize can be readily adapted to include covariates.)

Our specifications can also be modified to employ a variety of transformations of the threshold variables. Here, we focus on estimating only the marginal mean and variance of the implicit WTP variable. With normality, however, the admissible range contains the negative portion of the real line. Especially when regressors are employed, it will occasionally be the case that certain individuals will exhibit negative fitted values for their WTP. If it is deemed important to preclude negative fitted values, a logarithmic transformation of the thresholds, t_{1i} and t_{2i}, can be employed before the model is estimated. Likewise, the more general Box-Cox transformation is also viable and potentially very useful.

For our sample of 1013 respondents, Table 2 gives the results for a model without covariates under a range of different assumptions about β_1, β_2, σ_1, σ_2, and ρ. Model 1 is the most general, with all five parameters free to take on any value. The point estimate of the mean willingness to pay (WTP) for the first question is \$128.77, while the point estimate for the second question is \$146.06. The dispersion parameters for the fitted normal distributions are quite large, at \$339.51 for the first question and \$510.62 for the second. The error correlation is very precisely estimated at 0.9509.

Model 2 constrains the mean WTP underlying each response to be identical ($\beta_1 = \beta_2$). This "cross-equation" parameter restriction precludes estimation of this model using the packaged bivariate probit algorithms that could be used to produce the estimates shown for Model 1 (after an appropriate transformation). A likelihood ratio test with a value of only 0.48 indicates that this single parameter restriction cannot be rejected.

Model 3 allows the mean WTP to vary across the two responses, but constrains the dispersion parameter to be identical. The decrease in the likelihood is only 0.06, indicating that this single restriction cannot be rejected.

Model 4 constrains both the means and the dispersion parameters to be identical ($\beta_1 = \beta_2$, $\sigma_1 = \sigma_2$). Compared to Model 2, this single additional restriction is rejected by the likelihood ratio test statistic (which takes a value of 5.46, exceeding the critical value of 3.84). While the mean values elicited by the two questions appear to be statistically indistinguishable, there is significantly greater "noise" in the second-round responses. Compared to Model 3, however, the value of the likelihood ratio test statistic for the single additional restriction is only 3.82, falling just short of the 5% critical value of 3.84.

While restricting the σ's to be identical if the β's are constrained to be the same is rejected, it is nevertheless the case that *jointly* restricting both the means and the variances to be identical cannot be rejected. Model 4 involved two parameter restrictions compared to Model 1, and the log-likelihood decreases by only 2.97. Thus the likelihood ratio test statistic takes on a value of roughly 5.94, which falls just short of the 5% critical value for a $\chi^2(2)$ distribution: 5.99. Thus Model 4 could be argued to be the preferred specification.

5. Estimates under the Standard Assumptions

The last two columns of Table 2 illustrate the effects of alternative restrictions commonly employed in practice. Model 5 limits the analysis to simply the first question, where the offered threshold is truly exogenously determined. The point estimates of both β_1 and σ_1 are very similar to those pertaining to the first valuation question in the specification in Model 1.

Ignoring the second question is a sure way to avoid the problem of error correlations across the two questions. However, most previous studies explicitly using dichotomous choice with follow-up CV questions have assumed that $y_{1i} = y_{2i} = y_i$, so that $x_{1i} = x_{2i} = x_i$, $\beta_1 = \beta_2$, $\sigma_1 = \sigma_2$, and $\rho = 1$. The likelihood for the general model in (2) above becomes is undefined as ρ goes to 1, so we must construct a new likelihood function appropriate to this limiting case. If two normal distributions have the same mean and standard deviation, as their correlation goes to one, they become the identical variable. For each respondent, then, the two offered thresholds, t_{1i} and t_{2i}, then serve to divide the range of y_i into three regions. The two discrete responses, I_{1i} and I_{2i}, can be viewed as identifying which of these three regions contains the implicit valuation of the respondent.

Define R_{1i}, R_{2i}, R_{3i}, T_{li} and T_{ui} as follows:

(3) $R_{1i} = 1$ and $T_{ui} = t_{2i}$ if ($I_{1i} = 0$ and $I_{2i} = 0$);

$R_{2i} = 1$ and $T_{ui} = t_{2i}$, $T_{li} = t_{1i}$ if ($I_{1i} = 1$ and $I_{2i} = 0$), or

$R_{2i} = 1$ and $T_{ui} = t_{1i}$, $T_{li} = t_{2i}$ if ($I_{1i} = 0$ and $I_{2i} = 1$);

$R_{3i} = 1$ and $T_{li} = t_{2i}$ if ($I_{1i} = 1$ and $I_{2i} = 1$),

where each of the variables R_{ji}, j=1,2,3, is zero otherwise.

The likelihood function required to fit such a model can then be expressed as:

$$(4) \qquad \text{Log } L = \Sigma_i \{ \quad R_1 \log [\Phi((T_l - x'\beta)/\sigma)] + R_2 \log [\Phi((T_u - x'\beta)/\sigma) - \Phi((T_l - x'\beta)/\sigma)] + R_3 \log [1 - \Phi((T_u - x'\beta)/\sigma)] \}$$

Careful inspection of the likelihood employed in Hanneman, Loomis and Kanninen (1991) will reveal that this likelihood function is identical to theirs. The only difference is that two of their terms have been combined by the use of the R_2 indicator and the assignment of either the first or the second offered values as either T_u or T_l according to whether the response to the first question was "yes" or "no."

It may appear superficially that the $\rho = 1$ assumption with identical means and variances produces responses that can be analyzed like payment card interval data. If everyone received the identical two thresholds, packaged software for interval data could readily be employed (e.g. LIMDEP). Each pair of thresholds does indeed divide the range of the implicit underlying valuation variable. However, most packaged software programs require that everyone face the identical payment card. Here, there are many different "cards.",

In Table 2, the column for Model 6 shows the parameter estimates when the underlying valuations y_1 and y_2 are assumed to be identical. Clearly, this assumption vastly distorts both the implied mean value and the dispersion estimate relative to the less restrictive model with ρ unconstrained. It also causes a sharp decrease in the maximized value of the log-likelihood function. These results are striking because the assumption underlying them--that identical implicit valuations are being elicited by the two questions--would seem a perfectly plausible working hypothesis in many applications. Such an hypothesis is implicit in the analyses by Carson and Mitchell (1987) and by Hanemann, Loomis, and Kanninen (1991).

For this particular data set, it is *very* clear that the error correlation across the two implicit valuations is very strongly significantly different from both zero and one. For Model 1, the 95% confidence interval for our estimate of ρ is (0.9262, 0.9756). While highly correlated across the two questions, and possibly even drawn from the same *distribution* of values, the implicit valuations are not identical.

6. First Response Effects

It has occasionally been proposed that once a respondent has "made a commitment" by saying that they are willing to pay the first offered amount, they are more likely to say that they are also willing to pay the higher amount than they would be had they not received the first offer. On the other hand, it is also possible that respondents interpret the first offered amount as being the average social cost of the resource and balks at being asked whether they would be willing to pay "more than it costs."

If the respondent says "no" to the first offered amount, they may feel guilty about their unwillingness to pay that amount and be more likely to say yes to the smaller amount. Or, they may become annoyed if they perceive that the interviewer is trying to eke at least some money out of them by lowering the bid. They may say no to a lower second bid even though they might have said yes to that amount if it had been the only offer.

It is not possible with this data set to distinguish which of these behaviors might be occurring for each respondent. It is possible, however, to determine which effects might dominate by examining the effect on the mean and standard deviation of the WTP elicited by the second bid according to whether or not the response to the first offered amount was "yes" or "no." Table 3 displays these models. We examine three different specifications. Controlling for the correlation between the first and second responses, Model 7 allows mean WTP implied by the second question to vary according to whether the response to the first question was "yes" or "no" but restricts the standard deviation of WTP from the second question to be the same, regardless of the first response. The fitted mean WTP based on the second response for respondents who said "yes" to the first valuation question was, on average, $56 more than for respondents who said "no" to the first question, and this difference was statistically significant according to the coefficient's asymptotic t-test statistic. However, a likelihood ratio test of Model 7 against Model 1 reveals no improvement in the log-likelihood function.

Model 8 constrains the means elicited by the second question to be the same for the groups who respond "yes" and "no" to the first offer, but allows the standard deviations for the second question to differ. This model implies that the dispersion of valuations elicited on the second response is higher by $21 for people who responded "yes" to the first question. While the point estimate on this parameter is statistically significant according to its asymptotic t-test statistic, the maximized value of the log-likelihood function increases only by 0.03.

Model 9 allows both the mean and the dispersion for the WTP elicited by the first response to vary according to whether the first response was "yes" or "no." While the point estimates imply that both the mean and the dispersion are larger for the "yes" group, neither point estimate is significant and the log-likelihood remains statistically no higher than for Model 1.

The results for Models 7 through 9 suggest that the dominant effect is one wherein respondents who say "yes" initially are inclined to persist in saying "yes," even to higher amounts, and respondents who say "no" initially are inclined to persist in saying "no." The low relative frequency of yes/no and no/yes responses in these data corroborates this finding.

Models 10, 11, and 12 involve the same generalizations for first response effects as appear in Models 7, 8, and 9, with perhaps initially perplexing consequences. A dummy variable for a "yes" response to the first question is allowed to shift both the mean and the dispersion of the common distribution underlying both the first and the second questions, with bizarre effects on the other parameters. Why these consequences? These are inherently single-equation models and the dummy variable for a "yes" response to the first question is thoroughly endogenous. By putting an indicator for the magnitude of the valuation underlying the first response on the right-hand-side of this specification (without simultaneously modelling the first response and the correlation between the two), it becomes markedly easier to predict responses (as evidenced by the effect on the maximized value of the log-likelihood). However, due to the correlation between the new dummy variable and the implicit residual in this single equation model, endogeneity bias distorts the estimated parameters dramatically. First-response effect cannot be assessed in a single-equation setting such as this.

7. Starting Point Effects

It is also interesting to explore the sources of the apparent distortion in the valuation distribution point estimates due to Model 6 (with $\rho = 1$ imposed). We suspect that the starting point (i.e. the first offered value) may be unduly influential when this model is assumed. In a more standard context, starting point effects can be interpreted as the consequences of suggested values acting as cues for respondents who have no implicit agenda and are motivated only to provide a "socially correct" answer to a survey question. To explore the influence of starting points in Model 6, as opposed to Models 1 through 4, we can include dummy variables in each specification for first bids of $20, $50, and $100. The omitted category will be the $5 first bid. Table 4 displays the results of these specifications. The footnotes to the table detail how the results for Model 13 show that the two sets of starting point dummies have no statistically significant joint effect upon the fitted means implied for either the initial or the follow-up valuation questions.

The findings for Model 14 are similar, although only one set of dummy variables is involved because $\beta_1 = \beta_2$ in this specification. While none of the mean-shifting terms is statistically significant in these models, there is weak evidence in the point values suggesting that the implied mean of the value distribution may vary inversely with the magnitude of the starting bid.

Model 15 mimics Model 13 but constrains the σ parameters to be identical across responses. Again, none of the mean-shifting starting point dummy variables bears an individually statistically significant coefficient.

Only in Model 16 (which is the starting point generalization of our preferred Model 4) do we encounter one individually statistically significant starting point coefficient. However, the difference in the log-likelihood function achieved by freeing up these three additional parameters is only 0.56, so starting point effects are not significant overall in our specification.

In contrast, Model 17 demonstrates that starting point effects are definitely present under the standard assumptions of $\beta_1 = \beta_2$, $\sigma_1 = \sigma_2$, and $\rho = 1$. The fitted mean WTP decreases from $745 to $584 to $430 to $256 as the first offered value increases from $5 to $20 to $50 to $100 respectively. These differences are surprisingly large and the null hypothesis of no difference in mean WTP across starting points is firmly rejected by a likelihood ratio test. The conventional assumptions would appear to greatly exacerbate the researcher's perception of starting point distortions in respondents' answers.

8. Implications of Estimated Models

We have restricted the analysis in this paper merely to determination of the location (β) and scale (σ) of the implicit univariate marginal distributions of variables we describe as WTP at the instant of the initial contingent valuation question (y_1), and WTP at the instant of the follow-up question (y_2). We have assumed normality for the distributions of the true underlying values. For this particular data set, we are unable to reject the hypothesis that the mean WTP elicited by the first and second questions is the same. We can reject only the *incremental* assumption that the variances are identical (there appears to be greater dispersion in WTP values for the second question), but not the *joint* restriction that both means and variances are identical across the two questions. If the variance is indeed larger for the second question, it would seem natural to attribute this greater dispersion to the possibility that at least some portion of the sample is disconcerted by the follow-up question, or that opportunities for strategic responses have been perceived.

Model 4 appears to be preferred among the simple specifications without regressors examined in this exercise. With the large estimated value for the dispersion parameter, however, it is clear that many negative values of WTP are implied. While negative values of WTP might actually exist for some members of the population, policy makers may wish to place a minimum bound of zero on WTP in calculating a "mean WTP" for the environmental good in question. Whether such a minimum bound is seen as appropriate depends on whether WTP to support political beliefs such as "restrictions on the mining industry are undesirable" may properly be included in benefit-cost analysis. Some of the relevant issues are debated by Rosenthal and Nelson (1992), Kopp (1992) and Quiggin (1993).

We can take models where the fitted normal distributions imply some negative values and calculate the revised marginal mean if implied negative values are arbitrarily converted to zeros. This requires formulas for the expected value of a truncated normal distribution. If X is N(0,1), and we limit the domain of X to $X \geq c$, then E(X) is $\phi(c)/(1-\Phi(c))$. From Model 4, the mean of WTP over *only* the group with positive values is $437.02. However, the location and scale parameters imply that 37.52% of WTP values are negative. If the values for the group with negative WTP are set equal to zero, the overall mean of WTP from this question (across the two groups) is $271.28.

In contrast, for Model 6 (which we reject) the implied mean of WTP over only the group with positive values is $1353.51. The location and scale parameters imply that 36.14% of WTP values are less than zero. Setting the values for this group to zero and calculating the overall mean of WTP from Model 6 yields an average value estimate of

\$864.38. The imposition of $\rho = 1$ clearly produces a large upward distortion in the WTP estimates in this example.

9. Conclusions

It is critically important when analyzing responses from a dichotomous choice with follow-up contingent valuation survey to acknowledge the imperfect correlation between the responses to the first and second valuation questions. This simple illustrative example has highlighted the fact that serious distortions can potentially be introduced into valuation estimates by erroneously assuming that exactly the same implicit value underlies the respondent's reaction to each question (i.e. by constraining the distributional parameters to be identical and the correlation to be exactly unity). Furthermore, assuming that the implicit underlying "true" valuation is unchanged across responses precludes any assessment of first response effects. This assumption also appears to exacerbate starting point effects in our example.

The implication of the empirical findings in this paper is that respondents seem not to hold in their heads a single immutable "true" point valuation for an environmental resource. At best, they may hold a distribution of values--amounts they would be willing to pay with some associated probability density. This might be interpreted as "uncertainty." Whenever they are asked to produce a value for the resource, they make a draw from this distribution and use it as a basis for their response to the current discrete choice CV question.

Some inroads have been made recently in generalizing the usual contingent valuation micro-econometric theory to accommodate respondent uncertainty regarding their true values. Li and Mattsson (1993) have proposed a model that assumes respondents have incomplete knowledge about their true valuation of a non-market resource and thus may give incorrect yes/no responses regarding whether they would pay a given threshold amount for a resource. Their specification relies on debriefing information regarding respondents' subjective sureness of their responses.

In this paper, we have not entertained more-general models which introduce the complexity of additional covariates. For a truly random sample, covariates are not required to estimate the sample mean willingness to pay. When covariates are available, they frequently contribute substantially to the explanation of systematic variation in fitted valuations across individuals. Covariates can be especially useful because their presence reduces the *conditional* dispersion of the unobserved true valuation (σ in our specifications). All of our specifications can readily be generalized to include covariates.

While our example is only illustrative, it succinctly demonstrates how researchers can examine the stability of the implicit valuation distribution across the two stages of questioning in a dichotomous choice with follow-up survey. In other studies, the results will vary with the precise wording of the questionnaire, and with myriad other factors. We find in this case that *mean* WTP does not appear to vary significantly across the two questions, but the *dispersion* of valuations may be larger for the second question. (The incremental hypothesis tests concerning identical variances, given identical means, were borderline.) The implicit underlying point valuations are definitely not identical across the two questions, but fortunately, they are highly correlated (about .95). A joint hypothesis test of mean and variance suggests that the two values elicited *may* be drawn from the same distribution, but they are not the same number.

We hypothesize that the difference in the underlying point valuations between the first and second valuation questions may be artifact of some sort of strategic behavior, but its sources should most certainly be the subject of future inquiry. In any event, our findings are troublesome for the usual assumption that the two rounds of questioning should elicit (and therefore be able to bound) the same underlying point valuation. There is definitely some movement between the first and second questions in a dichotomous choice with follow-up survey. Further research on this subject is clearly necessary. It seems clear that the conventional assuredness of researchers that there exists some stable, fixed, "true" underlying valuation that is straightforwardly elicited by contingent valuation techniques deserves serious reconsideration.

REFERENCES

ABARE (1990) "Mining and the Environment: Resource use in the Kakadu Conservation Zone," AGPS, Canberra, Australia.

ABARE (1991) "Valuing Conservation in the Kakadu Conservation Zone," AGPS, Canberra, Australia.

Arrow, Kenneth, Robert Solow, Edward Leamer, Paul Portney, Roy Radner, and Howard Schuman (1992) *Report of the NOAA Panel on Contingent Valuation*, January 12, 1993.

Cameron, T.A. (1988) "A New Paradigm for Valuing Non-market Goods Using Referendum Data: Maximum Likelihood Estimation by Censored Logistic Regression", *Journal of Environmental Economics and Management*, **15**, 355-379.

Cameron, T.A. and D.D. Huppert (1991) "Referendum Contingent Valuation Estimates: Sensitivity to the Assignment of Offered Values," *Journal of the American Statistical Association*, 86, 910-918.

Cameron, T.A. and M.D. James (1987) "Efficient Estimation Methods for Closed-Ended Contingent Valuation Survey Data," *Review of Economics and Statistics*, 69, 269-276.

Carson, R.T., W.M. Hanemann, and R.C. Mitchell (1986) "Determining the Demand for Public Goods by Simulating Referendums at Different Tax Prices," manuscript, University of California, San Diego.

Carson, R.T., and R.C. Mitchell (1987) "Economic Value of Reliable Water Supplies for Residential Water Users in the State Water Project Service Area," report prepared for the Metropolitan Water District of Southern California.

Cummings, R.G., D.S. Brookshire, and W.D. Schulze (1986) *Valuing Environmental Goods: An Assessment of the Contingent Valuation*

Method, Totowa, NJ: Rowman and Allanheld.

Hanemann, W.M. (1991) Review appended to document: D. Imber, G. Stevenson, and L. Wilks (1991) "A Contingent Valuation Survey of the Kakadu Conservation Zone" RAC Research Paper No. 3, Resource Assessment Commission, Canberra, Australia.

Hanemann, W.M., J. Loomis, and B. Kanninen (1991) "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation," *American Journal of Agricultural Economics* 73, 1255-1263.

Imber, D., G. Stevenson and L. Wilks (1991) "A Contingent Valuation Survey Of The Kakadu Conservation Zone" RAC Research Paper No. 3, Resource Assessment Commission, Canberra, Australia.

Kopp, R. (1992), Existence Values Should Be Counted in Benefit Cost Analysis, *Journal of Policy Analysis and Management*, 11(1), 123-5.

Lehmann, E.L. (1983) *Theory of Point Estimation*, Wiley, New York.

Li, C.-Z. and L. Mattsson (1993) "Discrete Choice Under Preference Uncertainty: An Improved Structural Model for Contingent Valuation," Department of Forest Survey, Swedish University of Agricultural Sciences, Umea, Sweden.

Mitchell, R.C. and Richard T. Carson (1986) "Valuing Drinking Water Risk Reductions Using the Contingent Valuation Method: A Methodological Study of Risks from THM and Giardia," draft report to the U.S. Environmental Protection Agency, Washington, D.C.

Mitchell, R.C. and Richard T. Carson (1989) *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Washington, D.C.: Resources for the Future.

Patterson, D.A., and J.W. Duffield (1991) "Comment on Cameron's Censored Logistic Regression Model for Referendum Data," *Journal of Environmental Economics and Management*, 20, 275-283.

Quiggin, J. (1993), Existence Values and Benefit Cost Analysis: A Third View, *Journal of Policy Analysis and Management* (forthcoming).

Quiggin, J., R. Rose, and R. Chambers (1992) "Valuing Conservation in the Kakadu Conservation Zone (summary)," pp. 105-8 in Wallace, N. (ed.) *Natural Resource Management: An Economic Perspective*, ABARE, Canberra, Australia.

Rosenthal, D. and Nelson, R. (1992), Existence Values Should Be Counted in Benefit Cost Analysis, *Journal of Policy Analysis and Management*, 11(1), 116-22.

TABLE 1

Descriptive Statistics
(n = 1013)

Acronym	Description	Mean	Std. Dev.
t_1	Exogenous threshold for first question	$43.79	$36.29
t_2	Endogenous threshold for second question	$67.62	$77.85
I_1	Discrete response to first question (1 = yes, WTP amount; 0 = no, not WTP)	0.5892	
I_2	Discrete response to second question (1 = yes, WTP amount; 0 = no, not WTP)	0.5578	
	Joint frequencies of discrete responses:		
	$I_1 = 1$ and $I_2 = 1$	0.5084	
	$I_1 = 1$ and $I_2 = 0$	0.08983	
	$I_1 = 0$ and $I_2 = 0$	0.3520	
	$I_1 = 0$ and $I_2 = 1$	0.04936	

Thresholds: 1st(2nd)	n	Responses: YY NY	YN NN	%YY %NY	%YN %NN
$5(20/2)	253	150 7	17 79	59.3% 2.8%	6.7% 31.2%
$20(50/5)	252	136 11	20 85	54.0% 4.4%	7.9% 33.7%
$50(100/20)	255	124 15	23 93	48.6% 5.9%	9.0% 36.5%
$100(250/50)	253	105 17	31 100	41.5% 6.7%	12.3% 39.5%

Table 2

Estimation Results for General and Restricted Models (n = 1013)

Parameter	Model 1	Model 2 $\beta_1 = \beta_2$	Model 3 $\sigma_1 = \sigma_2$	Model 4 $\beta_1 = \beta_2$ $\sigma_1 = \sigma_2$	Model 5 just first response	Model 6 standard interval data model
β_1	$128.77 (6.88)	$144.82 (6.94)	$178.57 (7.28)	$150.22 (6.31)	$123.16 (4.08)	$510.10 (8.74)
β_2	$146.06 (7.08)	-	$150.61 (7.42)	-	-	-
σ_1	$339.51 (5.07)	403.42 (6.65)	$537.84 (6.41)	$472.42 (5.88)	$317.14 (2.87)	1437.81 (12.69)
σ_2	$510.62 (6.09)	523.72 (5.60)	-	-	-	-
ρ^a	0.9509 (77.12)	0.9493 (76.04)	0.9478 (78.49)	0.9520 (78.22)	-	1
Log L	-1080.86	-1081.10	-1081.92	-1083.83	-678.35[b]	-1364.55

[a] It may be useful in some applications to have the algorithm estimate $\rho^* = (1-\exp(-\rho))/(1+\exp(-\rho))$, in order to constrain the estimated value of ρ to lie strictly within the (-1,+1) interval. Here, we had no problem estimating ρ itself.

[b] This log-likelihood is not comparable to the others, since only the first valuation response is used in this model.

Table 3

First Response Effects
(n = 1013)

Parameter	Model 7[a]	Model 8[a]	Model 9[a]	Model 10[a]	Model 11[a]	Model 12[a]
β_1	$123.71 (5.76)	$128.63 (5.01)	$124.37 (4.74)	-$636.64 (-10.23)	$961.03 (19.31)	$-82.33 (-6.26)
β_2	$106.97 (4.09)	$144.29 (6.57)	$111.61 (4.21)	-	-	-
$\Delta\beta_2$ (yes_1)[b]	$ 56.49 (2.30)	-	$ 49.80 (0.97)	$1597.74 (15.46)	-	$71.27 (7.74)
σ_1	$319.08 (4.07)	$338.98 (3.77)	$321.67 (2.68)	$492.78 (13.62)	$6293.77 (29.22)	$1256.19 (14.56)
σ_2	$481.66 (4.47)	$491.06 (6.31)	$485.17 (4.22)	-	-	-
$\Delta\sigma_2$ (yes_1)[c]	-	$ 21.18 (2.60)	$ 0.03 (0.51)	-	-$5758.80 (-27.09)	$599.06 (9.47)
ρ	0.9251	0.9532	0.9284 (31.38)	1	1	1

	(49.71)	(78.00)				
Log L	-1080.86	-1080.83	-1080.83	-815.53	-1151.31	-750.86

[a] Models 7, 8, and 9 are two-equation models with the implicit values elicited from each referendum question treated as being jointly determined. Models 10, 11 and 12 are single-equation models that assume that both responses reflect the same underlying distribution of values, so that $\beta_1 = \beta_2$ and $\sigma_1 = \sigma_2$.

[b] This coefficient gives the shift in the mean implied by the second referendum question if the response to the first question was "yes." The base coefficient, β_2, is the mean implied by the second question if the response to the first question was "no."

[c] This coefficient gives the change in the standard deviation implied by the second referendum question if the response to the first question was "yes." The base coefficient, σ_2, is the standard deviation for the second question if the response to the first was "no."

TABLE 4

Starting Point Effects Added to Basic Models 1, 2, 3, 4 and 6

Coef.(first offer)	Model 13 (Model 1)	Model 14 (Model 2)	Model 15 (Model 3)	Model 16 (Model 4)	Model 17 (Model 6)
β_1 (t_1 = \$5)	\$133.90 (0.18)	\$187.60 (4.50)	\$228.39 (2.80)	\$188.11 (6.42)	\$745.28 (6.46)
$\Delta\beta_1$ (t_1 = \$20)	-20.54 (-0.10)	-35.37 (-0.56)	-46.59 (-0.73)	-35.75 (-1.21)	-161.39 (-1.05)
$\Delta\beta_1$ (t_1 = \$50)	-25.02 (-0.06)	-54.23 (-1.48)	-76.35 (-1.07)	-55.34 (-2.15)	-314.75 (-2.05)
$\Delta\beta_1$ (t_1 = \$100)	-44.22 (-0.01)	-45.82 (-1.50)	-77.30 (-0.94)	-38.90 (-1.33)	-488.71 (-3.15)
β_2 (t_1 = \$5)	\$178.29 (2.71)	-	\$178.29 (2.71)	-	-
$\Delta\beta_2$ (t_1 = \$20)	-32.71 (-0.53)	-	-32.71 (-0.53)	-	-
$\Delta\beta_2$ (t_1 = \$50)	-50.21	-	-50.21 (-0.75)	-	-
$\Delta\beta_2$ (t_1 = \$100)		-	-29.60 (-0.37)	-	-
σ_1			535.89 (3.30)		
σ_2			-		
ρ			0.9491 (61.95)		

	(-0.74) -29.61 (-0.37) 309.22 (0.17) 535.89 (3.31) 0.9491 (62.04)	437.53 (4.25) 565.38 (4.19) 0.9470 (71.42)		497.94 (4.82) - 0.9500 (72.66)	1427.91 (12.69) - 1
Log L	-1080.44[a]	-1080.48[b]	-1080.44	-1083.27[c]	-1358.78[d]

continued...

Table 4, continued:

[a] Maximized log-likelihood for Model 1 without starting point dummy variables was -1080.86. Chi-squared test statistic for the hypothesis that all six of the dummy coefficients are zero is only 0.84, so this hypothesis cannot be rejected.

[b] Maximized log-likelihood for Model 2 without these three starting point dummy variables was -1081.10. Chi-squared test statistic is only 1.24. Starting point effects are jointly insignificant.

[c] Maximized log-likelihood for Model 4 without these three starting point dummy variables was -1083.83. Chi-squared test statistic is only 1.12. Starting point effects are jointly insignificant.

[d] Maximized log-likelihood for Model 6 without starting point dummies was -1364.55. Chi-squared test statistic for restricting these three dummies to have zero coefficients is 11.54. Critical value is 7.81, so starting point effects are indeed jointly statistically significant in this model.

ENDNOTES

example, Hanemann, Loomis, and Kanninen (1991) employ models based upon logistic density functions. Consequently, their approach would not allow them to consider the presence of non-unitary correlations.

presence of regressors in a model would allow the researcher to explore any of a wide variety of alternative heoretic specifications for WTP, which might be interpreted as the equivalent variation associated with the spe hange in the resource. Linear-in-parameters models are the most popular (because they facilitate the use of pa lgorithms) but highly non-linear models are tractable when general function-optimizing algorithms are avail

w et al. (1992) Blue Ribbon Panel document details the "ideal" characteristics of a contingent valuation surve alue. To the extent that the Kakadu survey falls short of these ideals, the implied resource values will be called ion. However, we employ the Kakadu data primarily as a convenience sample to illustrate an important estima

as been some debate (Quiggin, Rose and Chambers, 1992, ABARE, 1991) over whether these responses should preted in WTP terms (implying a large proportion of respondents with WTP outside the range $0-$250) or in g' model, in which respondents with a precommitted position disregard the stated threshold, and answer "Yes" with their policy position.

nn, Loomis, and Kanninen (1991) utilize this assumption in their analysis of willingness to pay for wetlands in uin Valley of California. They use an alternative but equivalent formulation of the likelihood function that prese ble pairs of responses (YY,YN,NY,NN) for each observation. We opt first to identify the three *effective* interval dual. This highlights the correspondence between these models and payment-card interval data models with fix sholds across respondents.

sent, only the SURVIVAL subroutine in the SYSTAT package analysis package, and possibly a procedure in SAS, ap to analyze interval data in which the intervals are individual-specific. But general function-optimizing softw

ays be used.

eron and Huppert (1991), discrete-choice CVM with follow-up response data were simulated from actual paym

lustrate the potential efficiency gains from appending the follow-up question to a single discrete choice questio

ence of other distortions or biases. It was assumed that the true distribution of the unobserved valuation y was

ss all simulated formats: payment card, single dichotomous choice, and dichotomous choice with follow-up. T

cifications examined in this paper clearly require actual data.

Model 16, we restrict the starting point to have a *linear* effect on the fitted mean WTP, the coefficient on the in

rs an estimated coefficient of -0.39 (with a t-ratio of -1.37). Higher starting values have a small and insigni

n fitted mean WTP. Note that Model 13 cannot be converted to a model with a linear starting point effect, since

ficient on the starting point variable (in the model for the mean at the instant of the first question) would not b

npose upon Model 17 the assumption that the fitted WTP is linearly related to the starting value, the starting va

ficient of -4.84 (with a t-ratio of -3.22). On average, for each dollar higher is the initial offered value, the f

er by $4.84. This is a substantial and statistically significant effect.