Version: February 22, 1999

ALTERNATIVE NONMARKET VALUE-ELICITATION METHODS: Are the Underlying Preferences the Same?

by

Trudy Ann Cameron, UCLA
Gregory L. Poe, Cornell
Robert G. Ethier, Cornell
William D. Schulze, Cornell

ABSTRACT

In a unique survey, six different random sub-samples of respondents were presented with an opportunity to value the identical environmental good, each via a different elicitation method. The methods include: an actual purchase decision, a single-bid hypothetical dichotomous choice format with bids differing across respondents, an open-ended format, a payment card format, a five-level categorical scale for willingness to pay at each of 13 different monthly costs, and a stated choice among an extended set of five alternatives (including the reference good plus three other alternatives in conjunction with a "do-nothing" alternative). We employ a common underlying indirect utility function presumed to motivate utility-maximizing choices under *all* elicitation methods and a stochastic structure that is also fully compatible across methods. We specify a log-likelihood function that allows us to pool all of these different types of valuation data in one unified model. If we assume that preferences are identical across different types of people, this model allows us to demonstrate that the valuation results from the different elicitation methods are entirely compatible with the same underlying set of preferences, providing heteroscedastic errors across elicitation methods are permitted. If the preference functions are identical, then so too are the expected willingness to pay values. We find that identical preferences across elicitation methods cannot be rejected if errors are allowed to be heteroscedastic across elicitation methods. However, this tidy result seems to disappear if preferences are allowed to be heterogeneous across sociodemographic groups.

ALTERNATIVE NONMARKET VALUE-ELICITATION METHODS:
Are the Underlying Preferences the Same?

1. Introduction

Researchers who must resort to contingent choice survey methods to establish the social value of nonmarket environmental resources are by now familiar with the controversy over whether systematic biases in measured resource values can result from the investigator's choice of a value elicitation method. All of the commonly used elicitation methods are intended to elicit the same underlying preferences. If data collected using different formats lead to statistically indistinguishable estimates of the parameters of the assumed preference function, then it is said that "procedural invariance" holds (Tversky, Sattath, and Slovic, 1988; Kahneman and Tversky, 1984) and therefore that the different elicitation methods are equivalent. The existing literature offers several pair-wise comparisons of elicitation methods, but the evidence is fragmented.

In this paper, we employ a unique survey, designed specifically to allow comparison of values elicited by six alternative methods that have been employed elsewhere in the literature. By ensuring that everything about the survey instrument is strictly controlled across the different sub-samples involved and that the type of elicitation method is randomly assigned across respondents, we have an opportunity to directly compare the implied preference functions for respondents in each group. Unlike many past comparison studies, we can be confident that observed differences in values across methods do not stem instead from somewhat different descriptions of the environmental good, from different time frames for the survey, or from different populations being sampled.

Consider the situation where the researcher is willing to allow for differing error variances (heteroscedasticity across elicitation methods) in the stochastic portion of the indirect utility function that is assumed to drive individuals' responses to valuation questions. In this case, we show that it is not possible—for our data--to reject the hypothesis of identical underlying preference functions. This is an important result because if preferences are identical across value elicitation methods, then so will be expected willingness to pay for the good being valued.

In section 2, we situate this research in the context of prior literature on related issues. Section 3 describes our basic survey and its six variants, each involving a different method for eliciting non-market values of an environmental good. In section 4, we introduce our simple indirect utility-difference model and its corresponding willingness to pay function, compatible in both their systematic and stochastic components, that allow us to specify completely conformable empirical models for the choice data from each elicitation method. The empirical results are presented in Section 5. We describe the results from using each elicitation method in isolation, then pair-wise, and then when all of the data are pooled in one encompassing model. Section 6 considers the implications of our work for subsequent research efforts in this vein, and section 7 concludes.

2. Review of Prior Findings

a.) Comparing Different Elicitation Methods

Many previous studies have undertaken pair-wise comparisons between different elicitation methods. These studies have contributed to our conventional wisdom about the differences that can typically be expected among the values of non-market goods derived using these different methods. For example, the prevailing finding in comparisons of dichotomous choice (DC) contingent valuation methods with open-ended (OE) methods suggests that the dichotomous choice methods produce estimates that tend to be larger--sometimes much larger. A recent example is Loomis et al. (1997). Numerous comparisons are summarized in Schulze et al. (1996) and Brown et al. (1996). Some of the studies cited are Sellar et al. (1985), Boyle and Bishop (1988), Johnson et al. (1990), Walsh et al. (1992), Kealy and Turner (1993), and McFadden (1994). The DC/OE willingness to pay ratio generally seems to range between 1.1 and 5.

Other studies have compared dichotomous choice elicitation methods with payment card (PC) methods. These include Holmes and Kramer (1995), Ready et al. (1992) and Poe and Welsh (1996). The DC/PC willingness to pay ratio appears to range between 2.7 and 4.4 in these studies.

There are fewer instances of comparisons between multiple bounded (MB) elicitation methods and DC, PC, or OE methods. "Multiple bounded" is the term used here to describe an elicitation technique where the respondent is allowed to choose the *extent* to which he or she might be willing to pay. Sometimes, each respondent is also asked to rate their willingness to pay at each of several different "bid" values. Eliciting several choices from each respondent in this manner increases the available information about their preferences. Poe and Welsh (1996) have attempted some comparisons. The equivalence of MB and DC values is uncertain, but there is some evidence that MB values are moderately comparable to PC and OE value estimates.

Stated preference (SP) methods, now commonly associated with conjoint analysis methods originating in the marketing and transportation literatures (an referred to occasionally as “choice analysis” methods), ask respondents to choose among a set of scenarios that differ along several dimensions. Desvousges et al. (1983), Barrett et al. (1996), and Stevens et al. (1997) have explored these techniques in comparison with earlier CV formats and seem to find that SP value estimates exceed other types of CV estimates.

b.) Comparing Pairs of Elicitation Methods; Differing Variances

One significant antecedent for this paper is Boyle et al. (1996), which compares dichotomous choice with open-ended valuation responses, using independent samples and corresponding probit and tobit specifications allowing for differing error variances in the two sub-models. A common central tendency cannot be statistically rejected for two of their three pairs of data sets, but the estimated standard deviations are significantly different for all three pairs of data. Both the means and the standard deviations from the referendum-style samples exceed those from comparable open-ended data sets. These authors conclude that either open-ended questions underestimate values, or referendum-style questions overestimate them.

The present paper is differentiated from the Boyle et al. paper in that we employ the underlying utility-difference function as the basis from which *all* of our sub-model specifications are derived and we adopt the logistic error distribution assumption most common in random utility models. Furthermore, we combine not just pairs of samples, but six independent samples, all closely controlled except for their different elicitation methods.

c.) Comparing Stated and Revealed Preference Data with Different "Scales"

In the stated preference, or "conjoint analysis" literature, several papers have recognized the possibility of different scales of the latent variables underlying choices in different samples. Econometric tests of the equality of the scale parameters across samples have been conducted by Swait and Louviere (1993), by Adamowicz et al. (1994, 1997) and by Boxall et al. (1998).

The idea of different scales across real and hypothetical elicitation methods also appears in a comment by Haab et al. (1998) concerning a paper by Cummings et al. (1997), wherein the incentive compatibility of real and hypothetical responses is assessed. Haab and his collaborators undertake to re-estimate the Cummings et al.empirical model allowing for different error dispersions. They accomplish the estimation and testing of models (with and without heteroscedasticity) using packaged maximum likelihood discrete choice models. This is accomplished by conducting a line-search across values of the unknown proportionality parameter for the error variances in the two samples.

Our paper differs from papers in the conjoint tradition in that our basic log-likelihood function differs across our independent samples in far more ways than just the magnitude of the error dispersion parameter. In the conjoint tradition, the type of elicitation methods usually differ only in that some choices are observed and some are hypothetical although completely analogous. We have one observed discrete choice, one analogous hypothetical discrete choice, but also four additional, very different hypothetical choices. All of these must be accommodated within one unified specification for rigorous statistical testing of the equivalence of the underlying preferences.

Our paper also differs from the Haab et al. strategy in that we avoid the device of a grid search and instead specify and estimate a new (and tailor-made) utility function. With six pooled data sources, we normalize one disperson parameter to unity and estimate the other five as multiples of the first. These multiplicative factors are estimated jointly with the other model parameters, in a full information maximum likelihood procedure.

3. Description of Survey Versions

Our data consist of responses to both telephone and mail surveys, all of which were conducted with the cooperation of the Niagara Mohawk Power Company (NMPC) in Erie County of New York State. The topic of these surveys was the NMPC GreenChoice^{tm} program, wherein randomly selected households within NMPC's service territory were invited to consider either real or hypothetical additional charges on their utility bills in order to allow NMPC to plant trees and/or provide energy from renewable sources. Within the telephone mode and within the mail mode, the survey instrument was identical except for the manner in which consumer values for the proposed GreenChoice program were elicited. Elicitation methods were assigned randomly across households within each survey mode. Thus, there can be a presumption that the only explanation for systematic differences in the implied WTP within a survey mode is these different elicitation methods.

The data for the mail survey were collected from a random sample of households with listed telephone numbers from the NMPC service territory within Erie county. An advance notification letter was sent to candidate households on November 14, 1996, and an initial copy of the mail survey with a $2 incentive payment was mailed November 18. A first follow-up survey mailing to non-respondents took place on December 5, and a second follow-up on January 23, 1997.

The telephone survey targeted the same population. Advance notification letters with a $2 incentive were used here as well. Computer assisted telephone interviewing commenced June 6, 1996, and ran through August 11. At least eight contact attempts were made for households in the intended sample who could not be reached initially. Telephone interviews averaged 13 minutes in length. The earlier time frame for the telephone interviews (summer 1996 versus late fall and winter) means that it is conceivable that preferences may have shifted in the interim. Thus our actual choices form the mail survey may not be as compatible with the mail surveys as the different versions of the mail survey are among themselves.

Adjusted response rates to the different survey variants ranged from about 64% to 69% in the mail surveys and were about 70% in the telephone survey. See the Appendix and Table A1 for a description of non-response analysis and descriptive statistics on response proportions by survey variant.

The survey instrument was titled "Clean Energy and You." The preamble was common to all version of the survey. First, the GreenChoice^{tm} program was introduced as a voluntary partnership between Niagara Mohawk and its residential customers, designed to reduce air pollution and improve the environment in "our local communities." The program involves two parts: (i) using renewable energy, and (ii) planting trees. First, the distinction between non-renewable and renewable energy sources was described. Potential renewable energy sources were identified as wind, solar power, gas recovered from landfill sites. It was pointed out that these energy sources do not produce air or water pollution, they will conserve resources, but they also tend to cost more than other types of power.

The second part of the GreenChoice^{tm} program, if implemented, would plant thousands of trees on public lands throughout upstate New York. These planting projects would be developed with American Forests, the nation's oldest citizen conservation group. Respondents were then informed about the role of trees as "natural air filters, absorbing carbon dioxide (a contributor to global warming) and releasing oxygen into the atmosphere. When planted near buildings, trees help conserve energy by providing shade in summer and windbreaks in winter."

Following the description of the GreenChoice^{tm} program, a key component of this exercise was the description of how the program would work. The "provision point mechanism" and "refund" plan were described as follows:

> The Green Choice program would be funded voluntarily. Customers who decided to join the program would pay an additional fixed fee each month on their Niagara Mohawk bill. This fee would not be tax-deductible. Customers could sign up or cancel at any time. While customers sign up, Niagara Mohawk would ask for bids on renewable energy projects.
>
> Enough customers would have to become Green Choice partners to pay for the program. For example, if $864,000 were invested in the first year, it would allow Niagara Mohawk to plant 50,000 trees and fund a landfill gas project. The gas project could replace all fossil fuel electricity in 1,200 homes. However, if after one year participation were insufficient to fund green Choice activities, Niagara Mohawk would cancel the program and refund all the money that was collected.

The survey variants for this study were designed specifically to allow comparisons across elicitation methods with minimum ambiguity. In the sections below, we provide details on the elicitation format for the portion of the questionnaire that differs across subsamples.

Survey Version 1: *<u>(ACT) telephone actual purchase decision, $6 bid (Sample 1)</u>*

After the common preamble, respondents were posed the following question:

> So far I've described the GreenChoice program, as well as the $6 per month it would add to your household's electrical bill, if you were to join.
>
> You may need a moment to consider the next couple of questions. Given your household's income and expenses, I'd like you to think about whether or not you would be interested in the GreenChoice program.
>
> If you decide to sign up, we will send your name to Niagara Mohawk, and get you enrolled in the program. All your other answers to this survey will remain confidential.
>
> Does your household want to sign up for the program at a cost of $6.00 per month?

Survey Version 2: *<u>(MDC) dichotomous choice referendum, varying bids (Samples 2 through 8)</u>*

For this survey version, there were seven samples, each one with a different bid value. For samples

2 through 8, the bids were $0.50, $1, $2, $4, $6, $9, and $12, respectively. For each of these survey variants (and for Survey Versions 3 through 6 discussed below as well), the common preamble to the survey was followed by the budget constraint reminder: "Given your household's income and other expenses, we would like you to think about whether or not you would be interested in joining the GreenChoice program." The MDC valuation question was posed as follows (e.g. in the case of the $4 bid).

> 10. Would your household sign up for the program if it cost you $4 per month?
> *(Please circle ONE response)*
>
> 1 Yes
> 2 No

Survey Version 3: (OE) open-ended willingness to pay (Sample 9)

For this survey version, respondents were asked to indicate the highest amount they would willingly pay for the good. The common preamble was followed by:

> 10. What is the highest amount, if anything, that your household would pay each month and still sign up for the program? *(Please fill in amount below)*
>
> **$ __________ per month**

Survey Version 4: (PC) payment card (Sample 10)

For this survey version, respondents were offered a payment card and were asked to circle the highest amount willingly paid, from which we infer that their true willingness to pay is greater than or equal to this amount, but strictly less than the next highest amount shown on the card. We thus use the amounts to form intervals and situate each response in a unique interval.[1]

[1] There is some anecdotal evidence, from debriefing sessions with respondents who have faced payment cards, that respondents do not respond exactly to the instructions about the "highest amount" willingly paid. Instead, they may focus on the nearest amount. This interpretation is somewhat arbitrary, since we do not know who follows the instructions and who

10. What is the highest amount, if anything, that your household would pay each month and still sign up for the program?

(Please circle the HIGHEST amount you would pay PER MONTH for the program.)

0	$.50	$1	$1.50	$2
$3	$4	$5	$6	$9
$12	$16	$20	$25	$35
$45	$55	$75	$95	more than $95

Survey Version 5: *(MB) multiple bounded (Samples 11 through 13)*

For this survey version, respondents were given five possible responses to each of 13 different bids and asked to circle the response that best characterized their degree of willingness to pay each bid value. These 13 bids differed for samples 11, 12 and 13. In the first sample (Sample 11), the middle bid was $4. For this sample, the question format was as follows:

10. Would you joint the Green Choice program if it would cost you these amounts each month?

per month	Definitely No	Probably No	Not Sure	Probably Yes	Definitely Yes
10¢	A	B	C	D	E
50¢	A	B	C	D	E
$1	A	B	C	D	E
$1.50	A	B	C	D	E
$2	A	B	C	D	E
$3	A	B	C	D	E
$4	A	B	C	D	E
$6	A	B	C	D	E
$9	A	B	C	D	E
$12	A	B	C	D	E

adopts their own strategy.

	$20	A	B	C	D	E
	$45	A	B	C	D	E
	$95	A	B	C	D	E

In the second multiple-bounded sample (Sample 12), the middle bid was halved, to $2.00. The full range of bids was $0.10, $0.25, $0.50, $0.75, $1, $1.50, $2, $3, $4, $6, $12, $35, $95. For the third multiple-bounded sample (Sample 13), the middle bid was tripled, to $12.00. The individual bids were $0.10, $0.50, $1, $2, $4, $8, $12, $16, $20, $25, $35, $55, $95. Note that care was taken to ensure that the lowest and highest bids were the same across Samples 11 through 13, so that respondents would not be cued differently by the *range* of bids on these three instruments, only by differences in the distribution of bids.

Survey Version 6: *<u>(SP) stated preference (Samples 14 and 15).</u>*

In this version of the survey, respondents are invited to make choices between pairs of possible programs (of different complexity and at different costs). The presence of other program alternatives may potentially confound a respondent's choice between paying for the standard scenario (Option C) or not paying and not gaining this environmental enhancement (Option A). The different program scenarios for this version are:

Option A: pay nothing, get no environmental goods
Option B: plant 50,000 trees
Option C: plant 50,000 trees, provide renewable energy to 1,200 homes
Option D: plant 100,000 trees, provide renewable energy to 1,200 homes
Option E: plant 100,000 trees, provide renewable energy to 2,400 homes

For the stated choice surveys, each of Options B through E is first compared pair-wise with Option A, which amounts to a set of binary discrete-choice referenda for each respondent. Next, these same respondents are invited to choose the most-preferred of the full set of five options (which still includes the do-nothing option).[2] The numbers of trees and houses involved for each of Options A through E for Samples 14 and 15 are identical across respondents, but the prices for each option differ between these two samples. For Sample 14, the prices are $0, $0.50, $2.00, $2.50, and $4.00. For Sample 15, the prices are $0, $2.00, $6.00, $8.00, and $12.00, as depicted below.

There are different costs associated with each of the proposed options. Increasing the number of trees planted or the number of homes serviced by renewable energy sources raised the cost of the Green Choice program. Costs per household per month are given below for each of the five options discussed on the previous pages. For each option the program would remain voluntary. Note that Option A, no Green Choice program, would cost $0 per month. It is not presented below.

11. Option B, planting 50,000 trees, would cost $2 each month. Would you be willing to pay $2 per month for Option B? *(Please circle ONE response)*

1 Yes
2 No
3 Don't Know

12. Option C, planting 50,000 trees and providing renewable energy for 1,200 homes, would cost $6 per month. Would you be willing to pay $6 per month for Option C? (Please circle ONE response)

1 Yes
2 No
3 Don't Know

13. Option D, planting 100,000 trees and providing renewable energy for 1,200 homes, would cost $8 each month. Would you be willing to pay $8 per month for Option D?

1 Yes
2 No
3 Don't Know

14. Option E, planting 100,000 trees and providing renewable energy for 2,400 homes, would cost $12 each month. Would you be willing to pay $12 per month for Option E? *(Please circle ONE response)*

1 Yes
2 No
3 Don't Know

[2] Respondents were then invited to choose the next-most-preferred of these five options if their first choice was not available. However, respondents had difficulty with the idea that the "do nothing" option might not be available, so we elect not to assess whether these second-choices are consistent with the rest of the data.

Given your household's income and other expenses, we would like you to think about whether or not you would be interested in joining the Green Choice program, and if so, which option you would prefer. Below are a number of ways the Green Choice program could be implemented, including the cost of the program each month. You have the opportunity to choose your most preferred option.

15. If the Green Choice program was made available, which option would be your first choice? *(Circle the LETTER of your FIRST choice on the list below)*

Option	Cost per Month	Number of Homes Fueled with Renewable Energy	Number of Trees Planted
A	$0	0	0
B	$2	0	50,000
C	$6	1,200	50,000
D	$8	1,200	100,000
E	$12	2,400	100,000

4. Econometric Modelling of Choices

Whereas many early comparisons of WTP values produced by different elicitation methods employed entirely separate models for separate data sets, our objective in this paper is to create one "grand unifying model" that subsumes all the different types of choice data produced by our different survey variants. Only when a common preference structure and stochastic specification underlie all of the choice models is it possible to combine them all in a single model. In what follows, we will describe the components of this unified model individually. Since we employed a split sample design, we will then be able simply to add up the components in a single log-likelihood function to be maximized with respect to a common set of utility parameters that show up (differently) in each component. In this way, utility parameters and error distribution parameters can be restricted or unrestricted across elicitation modes, and likelihood ratio test statistics can be used to conduct formal hypothesis tests regarding these utility and error distribution parameters.

a.) The Common Preference Model

There are potentially five systematic varying parameters in the models we used to compare the indirect utility functions implied by respondents' choices under our six different elicitation methods. (These are versions 2 through 6 of the contingent choice mail survey, plus an analogous telephone survey requiring respondents actually to sign up to pay $6.) In the simplest case, each of these parameters can be just a constant, which assumes a common preference function for all respondents. If it is desired to allow preferences to vary systematically with demographic characteristics, for example, these constants can be generalized to functions of a set of explanatory variables (such as gender, age, etc.).

Recall that all five possible environmental enhancement scenarios are considered only in Version 6 of the survey, the stated preference (SP) variant. All of the other versions are concerned only with the choice between "doing nothing and paying nothing," (Option A) or choosing 50,000 trees and renewable energy for 1,200 houses at a price (Option C). Options B, D, and E (involving different numbers of trees and houses) appear only in Version 6.

The individual's level of indirect utility is presumed to depend upon the numbers of trees and houses affected, and the price of the proposed program. The four possible "do something" program alternatives can be captured by four dummy variables, B_i, C_i, D_i, and E_i, only one of which will be active for any program scenario. Indirect utility is then captured, in the simplest general case, by:

$$(4.1) \qquad V_i^1 = \beta_1{*}\, C_i + \beta_2{*}\, B_i + \beta_3{*}\, D_i + \beta_4{*}\, E_i - \beta_5{*}\, (Y_i - price_i) + u_i^1$$

We will assume that β_1* through β_4* are strictly positive, implying that the commodities being valued are "goods," not "bads." We will also assume that β_5* is negative, which is required for rationality, in the sense that indirect utility should not *increase* with price.[3] Consistent with the usual assumptions of random utility choice models, we will assume that the error term, u_i, has an extreme value distribution.

If Option A (the do-nothing alternative) is chosen, all of the dummy variables take on a value of zero as does $price_i$. Thus $V_i^0 = u_i^0$. In all but Survey Version 6, only Option C is being compared to Option A. For Survey Versions 1 through 5, then, we are estimating only β_1* and β_5* (or, in more general specifications, systematically varying versions of these parameters). The dummy variables $B_i = D_i = E_i = 0$ for all observations for these versions.

To prevent indirect utility from increasing with price, then, we restrict β_5* to be strictly negative by estimating β_5 * as -exp(β_5), where β_5 is a (potentially systematically varying) parameter which can take on any value dictated by the data. This restriction is empirically necessary primarily to ensure the theoretical plausibility of the models corresponding to Versions 3 and 4 of the survey (open-ended willingness to pay, and payment card models). We can similarly restrict the other ("intercept") parameters of the indirect utility function to be positive by estimating each β_j* as exp(β_j), j = 1,..,4. Then the underlying parameter β_j can take on any value. (Again, this is particularly important if β_j is to be converted to a systematic varying parameter that depends on the observed levels of respondent attributes. In counterfactual simulations, we do not wish to find predicted values of β_j* that are less than zero. While the indirect utility gleaned from the various combinations of the two environmental goods (trees and houses) may be very small (or even essentially zero), we choose to preclude the possibility that it might be negative.)

At least three different alternatives can be considered for parameterization of the indirect utility-difference function that drives individual choices: scalar unconstrained parameters, scalar parameters constrained to be positive and negative respectively, and systematic varying parameters constrained to be positive and negative, respectively. We illustrate for an individual's choice between Option C and the do-nothing alternative Option A:

---

[3] The sign convention in this equation may seem awkward, but it is designed to minimize errors of interpretation in the

(4.2) $$(V^1 - V^0)_i = \beta_1{*} + \beta_5{*}\ price_i + e_i\ , \text{ or}$$

$$(V^1 - V^0)_i = \exp(\beta_1) - \exp(\beta_5)\ price_i + e_i\ , \text{ or}$$

$$(V^1 - V^0)_i = \exp(x_{1i}'\beta_1) - \exp(x_{5i}'\beta_5)\ price_i + e_i\ ,$$

where $e_i = u_i^1 - u_i^0$ is distributed logistic$(0,\kappa)$.

Vital to the pooled-data model in this paper is the correspondence between the indirect utility-difference function (which drives the discrete choices) and the continuous maximum willingness-to-pay (WTP) function (which forms the basis of the open-ended and payment card responses). If we take the last version of the utility difference in (4.2), set it equal to zero, and solve for $price_i$, this dollar value will represent the predicted maximum willingness to pay by an individual with characteristics vectors x_{1i} and x_{5i}. The formula for fitted willingness-to-pay for Option C can therefore be expressed quite simply as:

(4.3) $$WTP_i = price_i = \exp(x_{1i}'\beta_1)/\exp(x_{5i}'\beta_5) + e_i/\exp(x_{5i}'\beta_5).$$

The first term in this expression is constrained to be positive, but the transformed error term is a scaled version of the underlying logistic $(0,\kappa)$ error, for which the support is the entire real line. If x_{5i} contains any more than just a constant term, this error term will furthermore be heteroscedastic across individuals.

It is possible to interpret the fitted conditional distribution of WTP_i as the distribution of individual WTP values in the portion of the population represented by this observation. Thus, the fact that the error term is unbounded may influence our strategies for calculating predicted expected WTP for an individual with given values for x_{1i} and x_{5i}. That portion of the fitted conditional density in the negative domain could be changed to a point mass at zero before the expected value is calculated. (This is akin to the interpretation of expected values in standard Tobit models.)

---

upcoming utility-difference function.

Incidentally, one can also consider the implications of fitted models of this type for the derivatives of $E[WTP_i]$ with respect to individual characteristics. Note that even with heteroscedasticity, $E[e_i] = 0$, so $E[WTP_i] = \exp(x_{1i}'\beta_1)/\exp(x_{5i}'\beta_5)$ if we disregard the problem of positive density in the negative domain. In this case, the derivatives are easy to calculate. If some particular individual attribute x_j is an element of both x_{1i} and x_{5i}, with corresponding estimated coefficients β_{1j} and β_{5j}, then the derivative of fitted WTP_i with respect to x_j is:

$$(4.4) \qquad \partial E[WTP_i]/\partial x_j = (\beta_{1j} - \beta_{5j})\, E[WTP_i] \,.$$

$E[WTP_i]$ is guaranteed to be non-negative by the exponentiation of the systematic varying parameters, so the sign of $\partial WTP_i/\partial x_j$ is determined by the sign of the parameter difference $(\beta_{1j} - \beta_{5j})$. Point estimates of the individual parameters are produced in the estimation process. Estimates of their difference and the standard error of this difference must be constructed explicitly from the point estimates and the asymptotic variance-covariance matrix of these estimates. In models that involve systematically varying utility parameters, we will be reporting t-test statistics for these differences, as well as a point estimate of the derivative of WTP with respect to each variable, computed at the sample mean values of all of the x_1 and x_5 variables.[4]

If an individual attribute x_j appears in only one "index," (either β_{1j} or β_{5j} but not both), then the derivative of WTP with respect to this parameter will be either β_{1j} WTP_i or $-\beta_{5j}$ WTP_i respectively. In the first case, the individual parameter estimate is sufficient to test hypotheses about the sign of the derivative. In the second case, care must be taken to remember the sign change.[5]

[4] At this juncture, we have not yet programmed computations of the approximate standard error of this fitted WTP at the means of the data. There exists the usual option of Monte Carlo sampling from the assumed joint normal distribution of the estimated parameters. This would allow one to build up a sampling distribution of the results of the formula for fitted WTP. However, since larger and larger samples, will increase the probability of getting parameters such that the simulated denominator value gets close to zero, there is a resulting concern that the sample mean of this distribution will grow continuously with the number of Monte Carlo replications.

[5] Our specially designed software makes these calculations and reports them automatically so as to avoid human error

b.) Survey Version 1: (ACT) actual dichotomous choice referendum, $6 bid (Sample 1)

Each individual in this sample is presented with the same $6 bid value and invited actually to decide to pay or not pay this amount for the scenario described (Option C). Indirect utility if this option is chosen is given by V_i^1; if it is not selected, indirect utility is V_i^0:

$$(4.5) \qquad V_i^1 = \exp(x_{1i}'\beta_1)(1)+\exp(x_{2i}'\beta_2)(0)+\exp(x_{3i}'\beta_3)(0)+\exp(x_{4i}'\beta_4)(0)+\exp(x_{5i}'\beta_5)\ price_i+u_i^1$$

$$V_i^0 = \exp(x_{1i}'\beta_1)(0)+\exp(x_{2i}'\beta_2)(0)+\exp(x_{3i}'\beta_3)(0)+\exp(x_{4i}'\beta_4)(0)+\exp(x_{5i}'\beta_5)(0)+u_i^0$$

Since Options B, D, and E are not being considered, the indirect utility-*difference*, $V_i^1 - V_i^0$, which drives this pair-wise choice can be simplified considerably:

$$(4.6) \qquad (V^1 - V^0)_i = \exp(x_{1i}'\beta_1)+ \exp(x_{5i}'\beta_5)\ price_i + e_i = Z_{1i} + e_i$$

where $e_i = (u_i^1 - u_i^0)$, and the coefficients β_2 through β_4 do not appear. Since $price_i$ does not vary across this sample (it is $6 for everyone), β_1* and β_2* cannot be separately identified without additional information (such as that gleaned from Survey Version 2). If x_1 and x_5 each consist solely of an "intercept" term, all that can be estimated is the sample average value of $Z_{1i} = (\beta^*_1 + 6\beta^*_5) = b_1$. As usual for binary discrete choice models, the scale of measurement of this indirect utility difference must be standardized by the unobserved logistic error dispersion parameter, κ, so that the dispersion parameter of e_i/κ is unity. It is an entirely arbitrary selection to use this particular survey variant for normalization of the dispersion. Dispersion parameters for all other variants will be estimated *relative to* this one.

If a respondent indicates they are willing to pay the bid amount, then let I_{1i}=1, otherwise I_{1i}=0. The contribution of Survey Version 1 to the log-likelihood function is therefore:

n_1

to the extent possible.

$$\text{(4.7)} \qquad \text{Log } L_1 = \sum_{i=1} I_{1i} \log\{\exp(Z_{1i})/[1+\exp(Z_{1i})]\} + (1-I_{1i}) \log\{1/[1+\exp(Z_{1i})]\}$$

Note that Log L_1 is just the familiar binary logit discrete choice model.

If β_1 * and β_5* are constants, a single parameter $Z_{1i} = b_1$ can readily be estimated. However, if we are assuming systematically varying preferences with sign constraints such that $\beta_1^* = \exp(x_i'\beta_1)$ and $\beta_5^* = -\exp(x_i'\beta_5)$, it does not seem possible to collapse the parameter space in any simple way. Of course, if we render $\beta_1^* + 6\beta_5^*$ into a *single* systematically varying "index" by substituting $x_i'\gamma$, instead of using *two* indexes in the form $\exp(x_i'\beta_1) – 6 \exp(x_i'\beta_5)$, it is possible to determine whether the probability that a respondent is willing to pay the $6 bid differs systematically across individuals with different characteristics. This model, however, is not nested within any of the other models considered in this study.

c.) Survey Version 2: (MDC) hypothetical dichotomous choice, varying bids (Samples 2 through 8)

Each individual in this sample is presented with a <u>different</u> bid value and likewise invited to indicate whether they would be willing to pay this amount for Option C. The indirect utility with and without Option C in this case is identical to that in equations (4.5). The important difference is that price_i now varies across observations. The indirect utility-difference index Z_{1i} above is thus replaced by $Z_{2i} = \exp(x_i'\beta_1) - \exp(x_i'\beta_5)\, \text{price}_i$. Thus, the indirect utility-difference that drives this pairwise choice will be:

$$\text{(4.8)} \qquad (V^1 - V^0)_i = Z_{2i} + e_{2i}$$

where u_{2i} is distributed logistic$(0,\kappa_2)$, with κ_2 not necessarily identical to κ_1. The coefficients β^*_2 through β^*_4 do not appear. Now, since price_i *does* vary in this sample, β^*_1 and β^*_5 can be separately identified.

With I_{2i} defined analogously to I_{1i}, indicating that the respondent is willing to pay the offered price for the basic scenario, and $I_{2i} = 0$ if the respondent is unwilling to pay this amount, the contribution to the log-likelihood function for Survey Version 2 is:

n_2

(4.9) $$\text{Log } L_2 = \sum_{i=1} I_{2i} \log\{\exp(Z_{2i}/\kappa_2)/[1+\exp(Z_{2i}/\kappa_2)]\} + (1-I_{2i}) \log\{1/[1+\exp(Z_{2i}/\kappa_2)]\}$$

The log-likelihood function Log L_2 is again the familiar binary logit discrete choice log-likelihood, except for the κ_2 parameter which allows the dispersion to differ from that for Survey Version 1. (We normalize κ_1 to unity in our pooled-data models).

d.) Survey Version 3: open-ended willingness to pay (Sample 9)

This version of the survey invites respondents to directly state their willingness to pay for the offered scenario. If behavior is consistent with simple economic theory of utility maximization, and if our model is appropriate, this willingness to pay should be that dollar price which would make the respondent indifferent between paying the price and getting the offered scenario, or not paying and not getting the scenario.

The relevant indirect utility difference should be rendered zero by the open-ended willingness to pay amount, WTP_i. Thus we should have:

(4.10) $$(V^1 - V^0)_i = \exp(x_{1i}'\beta_1) - \exp(x_{5i}'\beta_5)\, price_i + e_{3i}$$

Solving for $WTP_i = price_i$ such that $(V^1 - V^0)_i = 0$ yields the expression given in equation (4.3) above.

(4.11) $$WTP_i = \exp(x_{1i}'\beta_1)/\exp(x_{5i}'\beta_5) + e_{3i}/\exp(x_{5i}'\beta_5).$$

For this sample, as for the MDC sample, it may be prudent to allow the error term e_{3i} to be distributed logistic$(0,\kappa_3)$, where κ_3 can differ from κ_1 and κ_2.

For the open-ended elicitation method, respondents presumably provide a value for WTP_i. The conditional expected value of this distribution is $\exp(x_{1i}'\beta_1)/\exp(x_{5i}'\beta_5)$ and the (heteroscedastic) dispersion parameter is $\kappa_3/\exp(x_{5i}'\beta_5)$. Define Z_{3i} fundamentally differently from Z_{2i}, by using it to denote the "standardized" value of WTP_i:

(4.12) $$Z_{3i} = [\ WTP_i - \exp(x_{1i}'\beta_1)/\exp(x_{5i}'\beta_5)\ ]\ /\ [\ \kappa_3/\exp(x_{5i}'\beta_5)\ ]$$

To simplify the notation in what follows, define $s_{3i} = \kappa_3/\exp(x_{5i}'\beta_5)$.

The assumption of logistically distributed regression errors can adapted to a Tobit-like regression-by-maximum-likelihood context. The reason a Tobit-like model is indicated is because there is likely to be some heaping of reported WTP_i at zero, since negative values are not intuitively acceptable. Define $POS_i = 1$ if a strictly positive value of WTP_i is reported for observation i. If a zero WTP_i is reported, then $POS_i = 0$. The contribution of the responses to Survey Version 3 to the log-likelihood function is:

(4.13) $$\text{Log } L_3 = \sum_{i=1}^{n_3} \begin{array}{l} POS_i\ \{Z_{3i} - \log(s_{3i}) - 2*\log[1 + \exp(Z_{3i})]\} \\ \quad [1 - POS_i]\ \log\{\exp[Z_{3i}/(1+ \exp(Z_{3i}))]\} \end{array}$$

This component of the overall log-likelihood function can be described as an analog to the familiar Tobit log-likelihood, just based on the logistic error distribution rather than the more common normal distribution.

e.) Survey Version 4: payment card (Sample 10)

This version of the survey generates interval data for the true but unobserved WTP_i value. As in the open-ended case, the latent variable we must model is WTP_i. It can be defined as in equation (4.7), but we will now substitute e_{4i}, distributed logistic$(0,\kappa_4)$, where $\kappa_4 \neq \kappa_j$, $j \neq 4$.

(4.14) $$Z_{4i} = [\ WTP_i - \exp(x_{1i}'\beta_1)/\exp(x_{5i}'\beta_5)\ ]\ /\ [\ \kappa_4/\exp(x_{5i}'\beta_5)\ ]$$

If we treat the true but unknown value of WTP_i as the conditional expected value of WTP_i, the log-likelihood function involves the difference in the cumulative densities between the standardized upper bound and the standardized lower bound. Let $Z_{ui} = (t_{ui} - Z_{4i})/\delta_4$ and let $Z_{li} = (t_{li} - Z_{4i})/\delta_4$. For the assumed underlying logistic density function, the cumulative densities are $P_{ui} = 1/[1+\exp(-(Z_{ui}))]$ and $P_{li} = 1/[1+\exp(-(Z_{li}))]$. Then the contribution to the log-likelihood of the observations from Survey Version 4 is:

$$\text{(4.15)} \qquad \text{Log } L_4 = \sum_{i=1}^{n_4} \log\{P_{ui} - P_{li}\}.$$

Note that if we are dealing with the lowest interval, we will use just $\log\{P_{ui}\}$; with the highest interval, we will use just $\log\{1-P_{li}\}$.

This component of the overall log-likelihood can thus be characterized as a variant of the common interval-data model, adapted to an underlying logistic error distribution instead of the usual normal distribution.

f.) Survey Version 5: multiple bounded (Samples 11 through 13)

Respondents receiving Survey Version 5 were presented with 13 different bid values and asked to indicate (in categories) how likely they would be to be willing to pay each of these amounts. Five response alternatives were offered. We will denote these by y=yes, h=high, m=medium, l=low, and n=no. The intuitive framework for analyzing these responses is a multi-category generalization of the binary discrete choice referendum that applies for Survey Versions 1 and 2.

The indirect utility difference associated with paying the price and obtaining the scenario described (versus not paying and not obtaining the scenario) is presumed to drive the ordered categorical response to each question on this survey. Again, let the relevant indirect utility difference be:

$$\text{(4.16)} \qquad (V^1 - V^0)_i = \exp(x_{1i}'\beta_1) - \exp(x_{5i}'\beta_5)\ price_i + e_{5i} = Z_{5i} + e_{5i}$$

Note that we again allow the conditional error variance again to differ from that for other version of the survey.

Let Y_i=1 if the respondent chooses the "Definitely Yes" (yes) response, zero otherwise. Let H_i=1 if the respondent chooses the "Probably Yes" (high) response, zero otherwise. The indicators M_i, L_i, and N_i are defined similarly for the "Not Sure" (medium), "Probably No" (low) and "Definitely No" (no) responses.

There are four standardized estimated thresholds for the thirteen questions on Survey Version 5. Generally, in ordered-logit models, researchers arbitrarily set one interval threshold to zero since the location and scale of the underlying propensity to be willing to pay variable are unknown. To be consistent with the binary logit models also used in this study, however, we expect the zero level of the latent propensity variable in this ordered logit to be somewhere in the middle. The thresholds between the five intervals (no, low, medium, high and yes) are denoted as α_0, α_1, α_2, and α_3, respectively. Like the slope parameters in the binary choice models, these thresholds are known only up to a scale factor consisting of the dispersion parameter of the error term in this particular submodel, namely κ_5. For a single sample, then, we are able to estimate α_0/κ_5, α_1/κ_5, α_2/κ_5, and α_3/κ_5.

Then we can define five probabilities, one associated with each of the five categories within which each individual may have responded.

(4.17)
$$PY_i = 1/[1+\exp(\alpha_3/\kappa_5 - Z_{5i})]$$
$$PH_i = \exp(\alpha_3/\kappa_5 - Z_{5i})/[1+\exp(\alpha_3/\kappa_5 - Z_{5i})] - \exp(\alpha_2/\kappa_5 - Z_{5i})/[1+\exp(\alpha_2/\kappa_5 - Z_{5i})]$$
$$PM_i = \exp(\alpha_2/\kappa_5 - Z_{5i})/[1+\exp(\alpha_2/\kappa_5 - Z_{5i})] - \exp(\alpha_1/\kappa_5 - Z_{5i})/[1+\exp(\alpha_1/\kappa_5 - Z_{5i})]$$
$$PL_i = \exp(\alpha_1/\kappa_5 - Z_{5i})/[1+\exp(\alpha_1/\kappa_5 - Z_{5i})] - \exp(Z_{5i})/[1+\exp(Z_{5i})]$$
$$PN_i = \exp(\alpha_0/\kappa_5 - Z_{5i})/[1+\exp(\alpha_0/\kappa_5 - Z_{5i})]$$

The contribution to the log-likelihood function of the set of thirteen categorical responses for complete observations[6] from Survey Version 5 can then be expressed as:

[6] However many completed responses are provided to these thirteen questions by respondents receiving Survey Version

(4.18)

$$\text{Log } L_5 = \sum_{i=1}^{n_5} (1/13)\, \Sigma_1 \{ Y_i \log[PY_i] + H_i \log[PH_i] + M_i \log[PM_i] + L_i \log[PL_i] + N_i \log[PN_i] \}$$

We constrain the vector of α parameters to be identical across all thirteen questions (or portion thereof answered by any respondent). Note that the differing prices appear in the Z_{5i} terms that enter into each of the probability expressions.[7]

g.) Survey Version 6: stated preference (Samples 14 and 15).

This version of the survey introduces, as distractors, other possible scenarios in addition to the single scenario that is offered to respondents in Versions 1 through 5 of the survey. Of interest now is the indirect utility derived from each of the available scenarios. There are two samples for this version. One subsample was posed a lower set of prices for the set of scenarios, and the other was posed a higher set of prices. We will discuss the low-price instrument first (sample 14).

Rather than just V^1 and V^0, as before, we will now have five different indirect utility levels. Let these be denoted V^a (denoted V^0 above), V^b, V^c (denoted V^1 above), V^d, and V^e. Indirect utility under each of these choices can then be expressed in general as:

(4.19) $$V^j_i = \beta\, B^j_i + \beta\, C^j_i + \beta\, D^j_i + \beta\, E^j_i + \beta\, \text{price}^j_i + u_i$$

and in particular as:

(4.20) $$V^A_i = \beta(0) + \beta(0) + \beta(0) + \beta(0) + \beta(\$0.00) + u^A_i = Z_{Ai} + u^A_i$$

5, we scale the total contribution to the log-likelihood so that each respondent has a weight equal to respondents receiving other versions of the survey.

[7] The specification used in this study allows the researcher to restrict the α_0, α_1, α_2, and α_3 parameters to be identical across all thirteen responses. Alternately, they can be freed to take on distinct values for each of the thirteen categorical responses and one can test whether they might be identical. As an additional modification, each α_j parameter can be specified as a linear function of question number.

$$V^B{}_i = \beta(0) + \beta(1) + \beta(0) + \beta(0) + \beta(\$0.50) + u^B{}_i = Z_{Bi} + u^B{}_i$$
$$V^C{}_i = \beta(1) + \beta(0) + \beta(0) + \beta(0) + \beta(\$2.00) + u^C{}_i = Z_{Ci} + u^C{}_i$$
$$V^D{}_i = \beta(0) + \beta(0) + \beta(1) + \beta(0) + \beta(\$2.50) + u^D{}_i = Z_{Di} + u^D{}_i$$
$$V^E{}_i = \beta(0) + \beta(0) + \beta(0) + \beta(1) + \beta(\$4.00) + u^E{}_i = Z_{Ei} + u^E{}_i.$$

Since sample 15 has higher prices ($0.00, $2.00, $6.00, $8.00, and $12.00), there is independent variation across the pooled samples 14 and 15 in the prices faced by respondents.

We will allow for a different (common) error dispersion parameter from those that apply to the other samples. It would be possible to specify different dispersion parameters for each of the four pairwise choices made by respondents between each of Options B through E and Option A. Due to insufficient numbers of observations, however, we elect to specify $\kappa_{BA} = \kappa_{CA} = \kappa_{DA} = \kappa_{EA}$. We will refer to this common dispersion parameter as κ_{6P}, where the P denotes "pairwise choices."

For this survey version, however, respondents were also asked to choose their most-preferred option from among Options A through E. This is a five-alternative multiple choice model that we will model as a multinomial logit choice. The error dispersion for this choice context could be different again, so we will allow for a separate parameter κ_{6J}, with the J subscript denoting the "joint choice."

For each of the pairwise choices of Options B through E against the do-nothing Option A, the probabilities will be:

(4.21)
$$P_{Bi} = \exp(Z_{Bi}/\kappa_6) / [\exp(Z_{Ai}/\kappa_6) + \exp(Z_{Bi}/\kappa_6)]$$
$$P_{Ci} = \exp(Z_{Ci}/\kappa_6) / [\exp(Z_{Ai}/\kappa_6) + \exp(Z_{Ci}/\kappa_6)]$$
$$P_{Di} = \exp(Z_{Di}/\kappa_6) / [\exp(Z_{Ai}/\kappa_6) + \exp(Z_{Di}/\kappa_6)]$$
$$P_{Ei} = \exp(Z_{Ei}/\kappa_6) / [\exp(Z_{Ai}/\kappa_6) + \exp(Z_{Ei}/\kappa_6)]$$

We need indicators for a respondents choice in each of these four pairwise comparisons. Let I_{Bi}=1 if scenario B is chosen over A by respondent i, I_{Bi} = 0 otherwise. Similarly for I_{Ci}, I_{Di} and I_{Ei}. For each alternative j = B, C, D, and E, the contribution to the log-likelihood function will be

n_6

$$\text{(4.22)} \qquad \text{Log } L_{6P} = \sum_{i=1} \quad \sum_{j=B,C,D,E} \{ \; I_{ji} \log(P_{ji}) + (1 - I_{ji}) \log(1 - P_{ji}) \; \}$$

The second type of choice question posed to respondents who received the stated preference (SP) survey instrument asked them to choose their most-preferred option from the set of options A through E. We will allow for a separate dispersion parameter for this choice as well, denoted $\kappa_{6'}$. Simplify the notation to follow by defining $SUMP_i$ as $\exp(Z_{Ai}/\kappa_{6J}) + \exp(Z_{Bi}/\kappa_{6J}) + \exp(Z_{Ci}/\kappa_{6J}) + \exp(Z_{Di}/\kappa_{6J}) + \exp(Z_{Ei}/\kappa_{6J})$. Then under the logistic model, the probabilities of choosing each specific alternative from this set of five are given by:

(4.23)

$$P_{A}'_{i} = \exp(Z_{Ai}/\kappa_{6J})/SUMP_i$$
$$P_{B}'_{i} = \exp(Z_{Bi}/\kappa_{6J})/SUMP_i$$
$$P_{C}'_{i} = \exp(Z_{Ci}/\kappa_{6J})/SUMP_i$$
$$P_{D}'_{i} = \exp(Z_{Di}/\kappa_{6J})/SUMP_i$$
$$P_{E}'_{i} = \exp(Z_{Ei}/\kappa_{6J})/SUMP_i$$

As indicator variables for this top choice among the five possibilities, define $I_{Ai}' = 1$ if scenario (A) is most-preferred, $I_{Ai}' = 0$ otherwise. Similarly, define I_{Bi}', I_{Ci}', I_{Di}', and I_{Ei}'. The contribution to the log-likelihood of this first-choice program from among the five alternatives is then:

$$\text{(4.24)} \qquad \text{Log } L_{6J} = \sum_{i=1}^{n_6} \quad I_{Ai}'\log(P_{Ai}') + I_{Bi}'\log(P_{Bi}') + I_{Ci}'\log(P_{Ci}') + I_{Di}'\log(P_{Di}') + I_{Ei}'\log(P_{Ei}')$$

It seems appropriate that a single respondent to Survey Version 6 should have only unit weight in determining the maximized value of the log-likelihood function using pooled data. Thus, the number of pieces of choice information extracted from each respondent is used to divide the total contribution of each respondent to the overall log-likelihood function. For complete responses, $\text{Log } L_6 = (\text{Log } L_{6P} + \text{Log } L_{6J})/6$.

h. The complete specification

Due to the independence of the six samples, the log-likelihood function for the pooled data is just the sum of the six component likelihood terms. The parameters are found by maximizing:

$$\text{Log } L = \text{Log } L_1 + \text{Log } L_2 + \text{Log } L_3 + \text{Log } L_4 + \text{Log } L_5 + \text{Log } L_6 \tag{4.25}$$

The same utility parameters appear in each of the six components, and can be restricted or left unrestricted across the six terms as desired. The κ dispersion factors can be restricted to unity (making all dispersions the same as the (unidentified) normalized error dispersion for the numeraire sample. Alternately, the κ parameters for each survey version can be freely estimated.

5. <u>Empirical Results</u>

There are two key results from our empirical analyses. First, if individual preferences are assumed to be *homogeneous*, in the sense that preference function parameters do not differ systematically across individuals according to their sociodemographic characteristics, then neither do preferences appear to differ across elicitation methods. But our second important finding is that if individual preferences are allowed to be *heterogeneous* (that is, to vary systematically with a selection of observed sociodemographic characteristics), then there appear also to be differences in preferences *across elicitation methods*. The results for homogeneous preferences will be discussed first, followed by the results for heterogeneous preferences.

a.) Homogeneous Preferences

i.) Independent Estimates based on Different Methods with Separate Samples

First, we can consider the consequences of using each of our individual samples, with its own elicitation method, in an entirely separate empirical model. The first column of Table 1 provides these estimates. Of course, the 1-ACT sample, by itself, cannot produce a point estimate of WTP because the bid value does not vary across respondents. But the fact that the three implicit parameters in this model cannot be separately identified does not prevent us from maximizing the log-likelihood associated with this specification and using this unconstrained model for comparison with the achievable log-likelihood values for other models. Note that all other elicitation methods do allow us to produce a point estimate of the expected willingness to pay produced by that elicitation method.

Our specifications for the indirect utility-difference (and thus the corresponding WTP function) for each sub-sample are entirely conformable, unlike the case for the underlying specifications in many earlier comparisons across elicitation methods. The functional form for the systematic portion of each model is consistent with the same underlying utility function and the stochastic structures assumed for maximum likelihood estimation also correspond exactly across methods.[8]

What we see in the first column of Table 1 is familiar to researchers who have compared alternative elicitation methods: the point estimates of WTP differ rather markedly across our different samples--an outcome (in this case) attributable exclusively to the different elicitation methods that were used, since all other features of the sampling frame, survey instrument, and survey timing were controlled as rigorously as possible.

ii.) Pairwise Comparisons across Individual Value-Elicitation Methods

Most previous studies which have drawn comparisons between alternative elicitation methods have considered different methods two at a time. Table 1 also reports results for all possible *pairs* of samples. In each comparison cell there are two sub-cells. First, we report key results when all parameters are constrained to be identical across the two samples. Then we report results when the utility parameters are

[8] Consequently, it is possible that entirely ad hoc functional forms for the systematic and stochastic portions of these models might yield more-similar point estimates of WTP across methods. Then, however, there would be no rigorous method for conducting statistical tests of the equivalence of preferences.

constrained but the error dispersions are allowed to differ. The reported results are (i) the maximized value of the log-likelihood function for the joint model, (ii) the fitted point estimate of WTP for Option C based on the joint model, (iii) the number of estimated parameters in the joint model, and (iv) the aggregate number of choice occasions used to provide these estimates. (Sometimes more than one choice per individual is observed and used, although the total weight on each individual remains unitary). The final item reported is the factor of proportionality for the dispersion parameter κ for the non-numeraire sample in each pair. If the log of this value (the statistical quantity actually estimated) is statistically different from zero at the 5% level, there is an asterisk associated with this point estimate. Note that if $\log(\kappa)=0$, then $\kappa=1$, so that there is no difference between the two samples in the scale of the dispersion parameter.

The most important question with respect to each pair-wise comparison concerns whether the utility parameters are statistically the same across elicitation methods. If the sum of the log-likelihood values for each sub-sample (estimated separately) exceeds by a sufficient amount the maximized log-likelihood for a joint model with all corresponding parameters constrained to be identical, then we can reject the parameter restrictions in the constrained model. However, we are watching for evidence that freeing up the error variances means that the assumption of common corresponding utility parameters cannot be rejected. When the maximized log-likelihood appears in **bold face**, the restrictions in the joint model *are* rejected at the 5% level; when the log-likelihood appears in *italics*, the restrictions of the joint model *cannot* be rejected at the 5% level. Table 1 reveals that there is no pair-wise comparison for which *identical utility parameters* can be rejected, providing the error dispersions are allowed to differ across the two samples.

Why are we interested in the consequences of pooled data with parameter restrictions? If the indirect utility function parameters are in fact *the same* across samples with different elicitation methods, then the implied willingness to pay is also identical across these samples. If we were to rely on separate-sample estimates of the utility parameters, and were to compute fitted willingness to pay for each of these different samples, we would get the apparently conflicting implications captured by the WTP estimates in the first column of Table 1. For Survey Version 1, where the threshold does not vary across respondents, it is not possible to recover a point estimate of WTP. However, for the remaining 5 samples, the estimates vary rather widely: $2.84, $1.30, $2.09, $3.08, and $3.45 for samples 2 through 6 respectively. These are the types of discrepancies that have led to criticism of the robustness of value estimates across alternative hypothetical valuation methods.

iii.) Pooled Model: All Samples

Table 2 reports two sets of results for cases where the data from all six survey versions are pooled. First, we estimate a common set of parameters for the indirect utility function that might be presumed to underlie all of these choices.

As in the pair-wise comparisons considered above, we wish to compare the indirect utility-difference function implied by each sample (used individually) with the indirect utility-difference function estimated when preferences are constrained to be identical across all samples. Under an assumption of homogeneous preferences across individuals within a sample, the sum of the separate log-likelihood functions for the independent samples reported in the first column of Table 1 is -2712.768. The corresponding joint model, where all utility and dispersion parameters are restricted to be the same, achieved a maximized log-likelihood of only -2784.699, which clearly rejects these combined restrictions. However, if the utility parameters *remain* restricted, but we allow the dispersion parameters to differ across samples, the log-likelihood climbs to -2714.733, despite the large number of cross-sample indirect utility parameter restrictions remaining. These restrictions therefore cannot be rejected. This finding is one key result in this paper.

Some of the other results in Table 2 deserve comment. The point estimates of the intercept parameters for Options B, D, and E are based solely upon the choices by the roughly 325 members of the “stated preference” sub-sample (Sample 6). The estimates of WTP for the B, D, and E options suggest that people may be MOST willing to pay for the smallest program (Option B) and least willing to pay for (Option E), the most extensive program. This would appear to be the reverse of the usual "scope" effect, where people are expected to be willing to pay more if they get more.

This apparently anomalous result can be explored by appealing to other data we collected concerning individual preferences. All respondents were asked a preliminary question concerning "how interested" they were in the goal of replacing fossil fuel energy with renewable energy sources, and the goal of planting trees on public lands in upstate New York. The scale ranged from 1 (not at all interested) through 10 (very interested). About 9% of the sample failed to respond to each of these questions. Only 45% reported a rating of 6 or more for the renewable energy question (and only 17% rated it a 10); 70% reported a rating of 6 or more for the tree-planting question (and 39% rated it a 10). Respondents were clearly much more enthusiastic about the tree-planting exercise than about the renewable energy issue.

Unlike the other sub-samples, additional information was gathered from the 325 respondents to the stated preference question involving all five program options. These individuals were asked to rate their satisfaction with each of the five options on a scale of 1 (not satisfied at all) to 10 (very satisfied) under the assumption that each could be implemented *at no cost to them*. The proportion of individuals recording a score of 6 or higher for each of the options is A:8%, B:24%, C:42%, D:66%, and E:78%. If the programs are free, more is clearly better. However, if the programs are costly, people would rather do something than nothing, but do not want to do anything more than the minimum. This certainly raises the issue of a "warm glow" effect.

The second horizontal panel of Table 2 gives the estimated ordered-logit threshold values, relevant for the multiple-bounded value information from Sample 5. Recall that an ordinary binary logit model is simply a special case of an ordered logit with only two categories of response. In the ordinary logit case, the threshold between the two categories (willing to pay, not willing to pay) is set arbitrarily to zero. The corresponding threshold for the five-category multiple-bounded data should fall somewhere in the middle interval. This middle interval is bounded by α_1 and α_2. We would therefore expect to see α_0 and α_1 negative, and α_2 and α_3 positive. When different dispersion parameters are allowed for each sample, this expectation appears to be borne out. It does not seem to hold in the more restrictive specification that is rejected by our data.

The third panel of Table 2 describes the estimated factors of proportionality in the error dispersions for each data type. Survey Version 1 (the actual purchase decision sample, at $6) is defined as the numeraire sample, with dispersion parameter κ_1 normalized to unity. Note that in order to ensure positive dispersions, we estimate the *logs* of the multiples of this dispersion factor for each sample. For ease of interpretation, Table 2 reports the corresponding *levels* of these estimated dispersion factors. However, the associated t-test statistics still refer to the logged parameters. The asymptotic t-ratios can be used to test the hypothesis that the logarithm of the relevant parameter is zero (or, equivalently, that the estimated factor of proportionality for the dispersion is one, so that the dispersion is the same as for the numeraire sample). Note that the conversion to levels accounts for the inconsistency in some of the signs on the parameter point estimates and the t-ratios.

The varying-bid dichotomous choice variant (Survey Version 2), displays an error dispersion parameter κ_2 that is about 1.7 times as large as the dispersion for Survey Version 1 (the actual choices). Contingent dichotomous choice valuation data appears to be noisier than data on actual choices.

Conventional wisdom among researchers who have worked extensively with open-ended and referendum data holds that open-ended value estimates should be less variable. Here, however, the error dispersion in the open-ended data (Sample 3) is 2.6 times larger than for the numeraire 1-ACT sub-sample. The larger variance may be accounted-for by the existence of a handful of large outliers among the point values for willingness to pay in this sample. It may be that a Tobit-type conditional density based upon a logistic distribution is not fully compatible with the true distribution of WTP values in this sample.

Alternately, these outliers may be just that--random, but influential, anomalies.[9]

For the payment card sample, however, the dispersion is insignificantly different from that in the Version 1 (numeraire) sample. For these data, then, payment card elicitation seems to involve no more noise than do the actual purchase decisions. This result may or may not be generalizable. However, dispersion in the multiple-bounded (MB) data (Sample 5) is 3.5 times that in Version 1.

The dispersion measures for the pairwise and joint stated preference samples are also larger than for the actual choice data, being 1.7 times as large in the pairwise case. This finding is satisfying in that one would expect this dispersion to be very much like that for the similar dichotomous-choice sample of Version 2. The joint stated preference sample, however, is 3.2 times as large as that for the actual choice data (and about twice that of the pairwise contingent choice data).[10]

The common indirect utility parameters estimated by the pooled-data model produces a single common point estimate of WTP for Option C, which is reported in the fourth panel of Table 2. The value for the model with differing dispersion parameters is $2.28, which is larger than the separate estimates for the open-ended and payment card methods, but less than the separate estimates for the referendum, multiple-bounded, and stated-preference samples.

For the fitted distributions of WTP for each sample, the expected values are the same but the dispersions are different. Different dispersion means that the proportion of the population that would be predicted to vote in favor of the policy at any particular cost will differ across elicitation methods. The last main panel of Table 2 shows the consequences of differing fitted dispersion for the predicted portion of the represented population that would be willing to pay at least $6 for Option 6. Despite identical mean WTP,

---

[9] In future research with these data, we plan to explore the prospect for more general distributions that allow for skewness. Needed is a distribution with the desirable properties of the logistic, but with at least one additional shape parameter. The logistic model conforms nicely with the concept of utility maximization, but there will be other, more general distributions that may prove useful.

[10] In this paper, we do not account for the simultaneity among the pairwise and joint choices for Survey Version 6, which are made by the same individuals.

these proportions are 23% for the actual (ACT) data, 33% for the varying referendum (MDC) data, 39% for the open-ended (OE) data,[11] 22% for the payment card (PC) data, 23% for the multiple-bounded (MB) data, 34% for the pairwise stated preference (SP_P) and 41% for the multiple-choice stated preference (SP_J) data.

b.) Heterogeneous Individual Preferences

Our standard survey instrument included questions concerning the usual sociodemographic variables, including age and gender. We also employ income bracket information to construct a variable called "income" which is intended solely to control crudely for socioeconomic status. It is not accurate enough to be considered as budget constraint information, which accounts for why our linear indirect utility function specifications in this paper allow income to drop out of the indirect utility-differences assumed to drive choices. We also make use of some neighborhood variables. These include the proportion of owner-occupied housing in the respondent's zip code and the proportion of urbanized area in the respondent's zip code. These 1990 Census variables have been merged into our data set by means of 5-digit zip codes. As usual, they are measured with a degree of error which will be greater, the larger and more heterogeneous are individual zip code areas. However, as in other empirical valuation research, these neighborhood variables are sometimes significant.

Table 3 shows results for a set of models that allow for systematically varying preference parameters across sociodemographic groups. Notice that we currently generalize *only* the β_1* and β_5* parameters. The β_2*, β_3*, and β_4* parameters are estimated based solely on the information from the Version 6-SP sub-sample, and the limited number of observations available does not appear to support much generalization, although further exploration may yield more insight.

[11] Note that a logistic distribution adapted to a Tobit-like specification may be inappropriate for these open-ended data. The outliers that produce a large estimated dispersion parameter also contribute to a large upper tail for the fitted distribution.

There are at least some variables that appear to have a systematic effect on the sizes of the β_1* and β_5* parameters (the intercept and the slope of the indirect utility-difference function in the choice regarding Option C versus Option A (the status quo). For example, β_1* appears to decline with age. In those sub-samples where the shifter variables on the bid coefficient have a statistically significant effect, the marginal indirect utility from income, β_5*, appears to decline with both age and income, but to be larger with the proportion of owner-occupied housing in the respondent's zip code. For the actual choice sub-sample, the male dummy variable has a significant negative effect on the size of β_5*. For each model, we have combined the relevant effects and derived point estimates of the derivatives of WTP for Option C with respect to each of these sociodemographic shift variables. The evidence suggests that willingness to pay for the program is increasing in income, decreasing in age, higher for males, and inversely related to the proportion of owner-occupied housing (although this variable may be correlated with degree of urbanization).

Of course, this preliminary specification, selected for illustration of our findings for systematically varying preference models, is just one of a large number of alternative possible specifications. To facilitate pooling, we have elected to retain common sets of explanatory variables across all data types. We observe that the coefficient on the PURBIN variable as a shifter of the intercept term does not achieve individual statistical significance in any individual or pooled-data model, so it can probably be dropped. Further specification searching may, of course, yield alternative plausible specifications.

As in Table 2, the differentials in the dispersion parameters across elicitation methods are statistically significantly different from zero for all except the payment card method There remains strong evidence of heteroscedasticity, as was apparent in the earlier models with homogeneous preferences.

Once again, the most important question is whether it is possible to discriminate, statistically, between the preference functions implied by the separate samples. The sum of the individual log-likelihood values for the six separate samples is –2640.277. For the pooled data and its associated parameter restrictions, the value is –2670.611. A likelihood ratio test of the restrictions imposed when the data are pooled achieves a value of 60.668, whereas the 5% critical value for a $\chi 2(40)$ distribution is 55.759. For these data, one would only fail to reject the hypothesis of identical preference parameters at the 2% level of significance. Thus, the satisfying result of statistically identical preferences across elicitation methods seems to break down when individual attributes are allowed to interact with elicitation methods in the estimation of indirect utility parameters.[12]

Which sociodemographic differences contribute to the distinction in preference functions across elicitation methods? It is difficult to tell, because our specifications are highly non-linear in parameters. In a linear model, one could simply scale the estimated indirect utility-difference parameters by the estimated error standard deviation to gain a rough idea of the comparability of the point estimates on the shift variables. Here, it is not so straight-forward. For all except the 1-ACT sample, it is probably most appropriate to compare the fitted derivatives of WTP with respect to each shift variable, evaluated at the overall data means for all relevant variables. No standard errors can easily be calculated for these point estimates, so comparisons of these values across methods is admittedly crude. The income derivatives seem roughly the same order of magnitude for the MDC, OE and SP samples, but not for the PC and MB samples. The age effect is an order of magnitude smaller for the MB sample than for the others. The gender effects were not individually significant for any of these samples, and possibly as a consequence, bear conflicting signs. The owner-occupancy effects have consistent signs, but are almost twice as large for the SP sample as for any other sample. The urbanization variable was not statistically significant for any model, and not surprisingly, has a larger apparent effect for the OE sample than for the others. The places to look for contributions to the rejection of common preferences are most likely the variables with individually statistically significant coefficients. The best candidates are therefore the age effect and the

[12] An extensive appendix is available, concerning the problem of systematic non-response to different types of elicitation methods by different sociodemographic groups as captured by zip code level Census data for the neighborhood characteristics of potential respondents.

income effect. In both cases, the MB sample appears to be the most different.[13]

6. Directions for Future Research

a.) Different Forms for the Indirect Utility Function

In this paper, we have used a very simple functional form for the indirect utility-difference function that is linear in the price of the program. More exotic functional forms could certainly be explored, subject to the requirement that the corresponding formula for maximum willingness to pay must have an empirically tractable functional form. The more complicated the indirect utility-difference, the more unwieldy will be the associated WTP formula, and one must bear in mind that the error term from the former must also be carried through the derivation.

b.) Beginning with the Empirical Open-Ended Distribution

In this paper, we lean heavily on the tradition of using logit-type models to analyze discrete-choice contingent valuation data. This logit choice reflects the underlying extreme value distribution that is consistent with the utility theory underlying random utility models. However, the assumption of logistic errors derives only from this theory and represents a maintained hypothesis. In this paper, propagating the logistic errors through to the Tobit-type regression model used for the open-ended data maintains the utility-theoretic functional form and the stochastic structure throughout the model, but does not provide a particularly satisfying fit for the open-ended data.

It is possible that we could make better use of the open-ended data. Specifically, the open-ended sub-sample is the only one which directly gives us an array of point values for WTP. Some of these are

[13] If the 5-MB sample is omitted from the pooled data model, the maximized value of the log-likelihood is -2118.921. The summed separate log-likelihood values equal –2093.035. The likelihood ratio test statistic is therefore equal to 51.772. The unrestricted models have 49 implicit parameters, whereas the pooled model without MB has 16 parameters. The $\chi^2(33)$ critical value is 47.400. Even without the least-conforming MB sub-sample, then, we still reject the equivalence of the preference parameters for the five different elicitation methods other than the MB sub-sample. Similarly, if the 1-ACT sample is omitted from the set, it is still possible to reject the equivalence of the five other sets of preference parameters.

large positive outliers, and these tend to drive up the implied error variance for this sample. It may be fruitful to "begin" with the open-ended data and to choose a functional form for the distribution of WTP values that is consistent with the observed marginal distribution of values for this sample. Perhaps one could then work backwards from this distribution to an alternative set of stochastic assumptions for the other elicitation modes.

c.) Explaining Error Dispersions

We are presently working on the task of attempting to explain the observed differences in the magnitudes of the error dispersion parameters across elicitation modes. The present analysis shows that the implied error dispersions vary systematically with modes, and since modes were varied randomly across respondents, we do not expect the different modes to be picking up systematic differences in respondent characteristics across samples. However, even within a mode, there may be heteroscedasticity in WTP across individuals according to their characteristics. Cameron and Englin (1997) demonstrated heteroscedasticity in dichotomous choice contingent valuation data with respect to the amount of past experience each individual had with the environmental good in question. If the data-generating process for the errors (as a *function* of observable individual characteristics) differs systematically across samples, this could account for the different observed dispersions.

d.) More-General Specifications for the Multiple-Bounded Data

The α_j thresholds can be constrained to be identical for all 13 questions posed to each individual, or be allowed to differ by question by question. For each sample, then, there are potentially 39 additional threshold parameters to estimate, in addition to the basic parameters for the indirect utility function. However, since the first and 13th bid values were the same for all three versions of the multiple bounded instrument, it is not possible to estimate distinct sets of three thresholds for either question 1 or question

13. Questions 11 and 12 also pose problems for separate estimation since insufficient numbers of respondents chose the "Probably Yes" or "Definitely Yes" responses for these high bid values. Analysis is limited to a search for systematic tendencies in questions 2 through 10.

The estimated thresholds in an ordered logit model are typically interpreted as levels of the latent propensity variable, scaled by the error standard deviation of the conditional distribution of the propensity variable. Each of the three samples in this survey version may exhibit a different standard deviation for its error term. Likewise, we might allow for each question of the thirteen questions to exhibit a distinct error standard deviation. This lets us check for respondent fatigue (perhaps) if precision declines systematically with consecutive questions. Alternately, if precision increases with questions, perhaps learning (or crystallization of values) is occurring and respondents are increasingly able to identify their latent values as they progress through this list of choice questions.

7. Conclusions

For almost two decades, researchers have been puzzled by discrepancies among the values of non-market environmental goods according to the type of elicitation methods used to assess these values. Many past comparisons have been hampered by the need to use samples collected at different points in time, or from different populations, or via survey instruments that differ in other ways besides just the elicitation method employed. Early comparisons involved simply an inspection of the different point estimates of value produced by different methods, without conformable models and a rigorous statistical assessment of whether the values are statistically significantly different.

Our approach differs from earlier ones in a number of important ways. First, our different elicitation methods are part of a careful experimental design across split samples. All other aspects of the survey are carefully controlled so that any differences in the valuation results across samples can be attributed *only* to the different elicitation methods used.

Second, rather than comparing the mean WTP values implied by the different samples with their varying "treatments," we focus our efforts on the estimation of the underlying preference function parameters. This emphasis upon characterizing the preference function allows us to specify functional forms and stochastic structures that are entirely compatible across all of the different types of consumer choice scenarios that correspond to each elicitation method. The key insight is that if two or more elicitation methods imply the identical preference function for individuals in each sample, then these individuals will also have identical WTP. This is because the formula for expected WTP is derived from the indirect utility-difference function that is the foundation of all of our models.

Third, we zero in on the issue of heteroscedasticity across methods. In the context of dichotomous choice methods of elicitation, we do not estimate the indirect utility-function parameters individually. They are estimable only up to a scale factor (which is the dispersion parameter for the underlying latent utility-difference conditional distribution). Comparisons of the estimated parameters across different samples are invalid if the scale factors (the dispersion parameters) are different in the two groups. We allow the our utility-differences for each elicitation mode to exhibit whatever degree of dispersion the data dictate (subject to normalization on any one arbitrarily chosen sample).

What do we find? The evidence confirms the growing concensus that it is essential to accommodate heteroscedasticity across elicitation methods. With otherwise homogeneous preferences across individuals, there appears to be heteroscedasticity across elicitation methods, and this heteroscedasticity seems more-or-less consistent with the findings in the stated preference (conjoint analysis) literature that different methods lead to different “scale factors” (the common term for error dispersions). Even with the more than 7000 choices that we employ to fit our most elaborate specification that pools the data for all types of elicitation methods, we are not able to reject the hypothesis of identical indirect utility-difference functions across elicitation methods (i) provided preferences are assumed to be common across sociodemographic groups and (ii) provided we allow for heteroscedasticity across methods. If we attempt to impose homoscedasticity, however, the restriction of identical preference parameters is rejected.

On the topic of heteroscedasticity, there have been a couple of attempts to explain apparent differences in the amount of random noise associated with responses to different elicitation methods. Harrison (1989) and Smith and Walker (1993a,b) have suggested, bases on their laboratory results, that valuations will be more variable the lower the opportunity cost to respondents of deviating from the rational decision. Smith and Walker (1993b) show that bid function slopes are insensitive to these opportunity costs, but response variability can be decreased markedly by imposing larger opportunity costs.

The initial finding of identical preference parameters, but heteroscedasticity across methods, might have provided a satisfying conclusion, had we not elected to dig further by introducing systematic varying parameter models. When preferences are allowed to vary more flexibly across individuals, a pooled model that restricts all of the systematic utility parameters to be identical across elicitation methods is rejected in favor of six independent models on the separate sub-samples. While it is possible that aggressive specification-searching across all possible subsets of elicitation methods my yet turn up cases where it is not possible to reject preference equivalence across elicitation methods, such a case has not yet been found among the systematic varying parameter specifications. Where we go from here, as researchers, is an open question.

REFERENCES

Adamowicz, W.L., J. Louviere, and M. Williams (1994) "Combining Stated and Revealed Preference Methods for Valuing Environmental Amenities," *Journal of Environmental Economics and Management* **26**, pp. 271-296.

Adamowicz, W.L., J. Swait, P.C. Boxall, J. Louviere, and M. Williams (1997) "Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation," *Journal of Environmental Economics and Management* **32**, pp. 65-84.

Boxall, Peter C., Jeffrey Englin, and Wiktor L. Adamowicz (1998) "Valuing Undiscovered Attributes: A Combined Revealed-Stated Preference Analysis of North American Aboriginal Artifacts," mimeo, Department of Applied Economics and Statistics, University of Nevada, Reno.

Boyle, Kevin J., and Richard C. Bishop (1988) "Welfare Measures Using Contingent Valuation: A Comparison of Techniques," *American Journal of Agricultural Economics* **70**(1), 20-28.

Boyle, Kevin J., F. Reed Johnson, Daniel W. McCollum, William H. Desvousges, Richard W. Dunford, and Sara P. Hudson (1996), "Valuing Public Goods: Discrete versus Continuous Contingent-Valuation Responses," *Land Economics* **72**(3), August, pp. 381-96.

Brown, Thomas C., Patricia A Champ, Richard C. Bishop, and Daniel W. McCollum (1996) "Which Response Format Reveals the Truth about Donations to a Public Good?" *Land Economics* **72**(2), May, pp. 152-66.

Barrett, C., T.H. Stevens, and C.E. Willis (1996) "A Comparison of CV and Conjoint Analysis in Groundwater Valuation," *W-133 Benefits and Costs Transfer in Natural Resource Planning*, 9, 79-104.

Cameron, Trudy Ann, and Jeffrey Englin (1997) "Respondent Experience and Contingent Valuation of Environmental Goods," *Journal of Environmental Economics and Management* **33**(3), 296-313.

Cameron, Trudy Ann, W. Douglass Shaw, and Shannon Ragland (1998) "Nonresponse Bias in Mail Survey Data: Salience vs. Endogenous Survey Complexity," in *Valuing the Environment Using Recreation Demand Models*, Joseph A. Herriges and Catherine L. Kling (eds.) Edward Elgar Publishing Ltd., forthcoming.

Cummings, Ronald G., et al. (1997), "Are Hypothetical Referenda Incentive Compatible?" *Journal of Political Economy* 105, 609-621.

Desvousges et al. (1993) "Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability," in Hausman, Jerry A. (ed.) *Contingent Valuation: A Critical Assessment*, New York: Elsevier Science, 91-159.

Harrison, Glenn (1989) "Theory and Misbehavior in First Price Auctions," *American Economic Review*, 749-62.

Holmes, Thomas P., and Randall A. Kramer (1995) "An Independent Sample Test of Yea-Saying and

Starting Point Bias in Dichotomous-Choice Contingent Valuation," *Journal of Environmental Economics and Management* **29**, 121-132.

Johnson, Rebecca L., N. Stewart Bregenzer, and Bo Shelby (1990) "Contingent Valuation Question Formats: Dichotomous Choice versus Open-Ended Responses," in Johnson, Rebecca L, and Gary V. Johnson (eds.) *Economic Valuation of Natural Resources: Issues, Theory, and Applications*. Social Behavior and Natural Resources Series, Boulder and Oxford: Westview Press, 193-203.

Kealy, Mary Jo, and Robert W. Turner (1993) "A Test of the Equality of Closed-Ended and Open-Ended Contingent Valuations," *American Journal of Agricultural Economics*, **75**(2) May, 321-31.

Kahneman, D. and A. Tversky (1984) "Choices, Values and Frames," *American Psychologist* **39**, 341-50.

Loomis, John, Thomas Brown, Beatrice Lucero, and George Peterson, "Evaluating the Validity of the Dichotomous Choice Question Format in Contingent Valuation," *Environmental and Resource Economics*; **10**(2), September 1997, 109-23.

Loomis, John B., Michael Lockwood, and Terry DeLacy (1993) "Some Empirical Evidence on Embedding Effects in Contingent Valuation of Forest Protection," *Journal of Environmental Economics and Management*, **25**(1), Part 1, July, 45-55.

Loomis, John B. (1990) "Comparative Reliability of the Dichotomous Choice and Open-Ended Contingent Valuation Techniques," Journal of Environmental Economics and Management, **18**(1), January, 78-85.

Mackenzie, John, "A Comparison of Contingent Preference Models," *American Journal of Agricultural Economics*; **75**(3), August 1993, 593-603.

McFadden, Daniel (1994) "Contingent Valuation and Social Choice," *American Journal of Agricultural Economics*, **76**(4:Nov), 689-708.

Milon, J. Walter (1989) "Contingent Valuation Experiments for Strategic Behavior," Journal of Environmental Economics and Management, 17(3), November, 293-308.

Poe, Gregory L. and Michael Welsh (1996) "Uncertainty in Contingent Values and Procedural Variance Across Elicitation Formats: A Multiple Bounded Discrete Choice Approach," Working Paper Series in Environmental and Resource Economics 96-03, Cornell University.

Ready, Richard C., J.C. Buzby, and D. Hu (1996) "Differences Between Continuous and Discrete Contingent Valuation Estimates," *Land Economics*, **72**(3:Aug), 397-411.

Schulze, William D., G. McClelland, D. Waldman, and J. Lazo (1996) "Sources of Bias in Contingent Valuation," in D.J. Bjornstad and J.R. Khan, eds., *The Contingent Valuation of Environmental Resources*, Edward Elgar Publishers, Cheltenham, UK.

Sellar, Christine, J.R. Stoll, and J.-P. Chavas (1985) "Validation of Empirical Measures of Welfare Change: A Comparison of Non-Market Techniques," *Land Economics* **40**(2:May), 156-175.

Smith, Vernon L., and J.M. Walker (1993a) "Rewards, Experience, and Decision Costs in First Price

Auctions," *Economic Inquiry* **31**, 237-245.

Smith, Vernon L., and J.M. Walker (1993b) "Monetary Rewards and Decision Cost in Experimental Economics," *Economic Inquiry* **31**, 245-261.

Stevens, Thomas H.; Barrett, Christopher; Willis, Cleve E. (1997) "Conjoint Analysis of Groundwater Protection Programs," *Agricultural and Resource Economics Review* **26**(2), October, 229-36.

Tversky, A., S. Sattath and P. Slovic (1988) "Contingent Weighting in Judgment and Choice," *Psychological Review* **95**, 371-384.

Table 1
Pairwise Comparisons of WTP Estimates by Different Elicitation Methods

	solo	2-MDC		3-OE		4-PC		5-MB		6-SP	
		same κ	diff. κ	same κ	diff. κ	same κ	diff. κ	same κ	diff. κ	same κ	diff. κ
1-ACT	-62.940 $ (n/a) np=1 n=116	**-520.614** $2.57 np=2 n=851	*-517.559* $2.84 np=3 κ_2=1.91*	**-772.424** $0.29 np=2 n=356	*-769.134* $1.31 np=3 κ_3=2.06	*-746.834* $2.15 np=2 n=387	*-746.718* $2.09 np=3 κ_4=0.91	**-621.291** $0.00 np=6 n=4956	*-617.165* $0.31 np=7 κ_5=2.30	*-315.947* $1.75 np=6 n=1373	*-313.953* $3.45 np=7 κ_{6P}=2.46 κ_{6J}=4.49*
2-MDC	-454.619 $2.84 np=2 n=735			**-1165.927** $2.01 np=2 n=975	*-1162.021* $2.54 np=3 κ_3=1.53*	**-1151.187** $2.67 np=2 n=1006	*-1139.362* $2.33 np=3 κ_4=0.55*	**-1017.392** $2.07 np=6 n=5575	*-1008.844* $2.84 np=7 κ_5=2.16*	*-705.746* $2.89 np=6 n=1992 κ_{6J}=1.89*	*-705.728* $2.90 κ_{6P}=1.05 κ_{6J}=1.94*
3-OE	-706.194 $1.30 np=2 n=240					**-1440.185** $1.85 np=2 n=511	*-1390.320* $2.01 np=3 κ_4=0.37*	**-1262.349** $1.11 np=6 n=5080	*-1260.419* $1.31 np=7 κ_5=1.35*	*-959.207* $1.76 np=6 n=1497 κ_{6J}=1.56	*-958.089* $1.93 np=7 κ_{6P}=0.68 κ_{6J}=1.28
4-PC	-683.778 $2.09 np=2 n=271							**-1287.151** $1.80 np=6 n=5111	*-1238.003* $2.09 np=7 κ_5=3.69*	**-939.761** $2.23 np=6 n=1528 κ_{6J}=	*-935.281* $2.16 np=7 κ_{6P}=1.84* κ_{6J}=3.43*
5-MB	-554.225 $3.08 np=6 n=4840									**-808.499** $3.78 np=10 n=6097	*-805.236* $3.45 np=11 κ_{6P}=0.48* κ_{6J}=0.88
6-SP	-251.012 $3.45 np=6 n=1257 κ_{6J}=1.83*										

KEY: Cells for "same κ" models contain maximized value of the log-likelihood function, fitted WTP for option C, number of estimated parameters, number of choices employed in model; Cells for "diff. κ" models contain maximized value of the log-likelihood, fitted WTP for option C, number of estimated parameters, point estimate of κ_j factor for sample j. Note that 6-SP data allows differing κ_{6J} for multiple choice (versus pairwise choices) in all models. Bold-face log-likelihood indicates restrictions (against separate models) are rejected at 5% level of significance; italics indicates failure to reject.

Table 2: Pooled Samples; Utility Parameters Constrained

Type of Parameter	Common κ across data types	Differing κs across data types
Utility parameters:		
$\exp(\beta_{11})$ * Option C_i +	-1.048 (-7.45)**	-0.310 (-1.40)
$\exp(\beta_{21})$ * Option B_i +	0.152 (0.69)	0.724 (2.32)**
$\exp(\beta_{31})$ * Option D_i +	-2.431 (-0.79)	-1.397 (-0.64)
$\exp(\beta_{41})$ * Option E_i -	-16.26 (-0.10)	-19.86 (-0.07)
$\exp(\beta_{51})$ * bid_i	-1.756 (-42.42)**	-1.136 (-68.88)**
Ordered logit thresholds:		
α_0	-1.228 (-11.16)**	-3.285 (-4.61)**
α_1	-0.6542 (-7.48)**	-1.481 (-3.37)**
α_2	-0.1525 (-2.39)**	0.09364 (0.25)
α_3	0.7542 (7.09)**	3.047 (3.72)**
κ multiples:[a]		
κ_1 (actual)	1.0	1.0
κ_2 (MDC)	1.0	1.732 (2.85)**
κ_3 (OE)	1.0	2.575 (5.22)**
κ_4 (PC)	1.0	0.9505 (-0.30)
κ_5 (MB)	1.0	3.514 (5.83)**
κ_6 ($SP_{pairwise}$)	1.0	1.735 (2.01)**
κ_7 (SP_{joint})	1.813 (2.16)**	3.236 (3.54)**
Fitted WTP at means of data		
Option C	$2.03	$2.28
Proportion WTP $6		
1 - ACT	33.5 %	23.3 %
2 - MDC	"	33.4 %
3 - OE	"	38.6 %
4 - PC	"	22.2 %
5 - MB	"	23.3 %
6 - $SP_{pairwise}$	"	33.5 %
7 - SP_{joint}	40.7 %	40.9 %
Number of choices employed	7459	7459
Maximized Log L	-2784.699 [c]	- 2714.733 [d]

[a] κ multiples are estimated as powers of e to ensure that they remain positive. For ease of interpretation, the point estimates have been exponentiated. The t-test statistics, however, remain tests of the hypotheses that the estimated exponents are zero. If the exponent is zero, the κ multiple is unity (exp(0)) and the dispersion for the sample in question is not different from that of the numeraire 1-ACT sample.

[b] Percentages differ only because conditional dispersion of WTP distribution differs across data types. E[WTP] is constrained to be identical for all samples. Logistic distributional assumption is strong (especially in 3 - OE Tobit-type specification).

[c] The sum of the maximized log-likelihood values for the separately estimated models (see Table 1) is -2712.768. There are 25 implicit parameters across the six differ data types, although only 21 of them can be identified. There are only ten parameters in this pooled-data model. The likelihood ratio test statistic for the restrictions embodied in this model is 143.8. Regardless of how one counts the number of restrictions they are clearly rejected.

[d] There are five fewer restrictions in this model, since the κ parameters for each data type are allowed to differ. The likelihood ratio test statistic for the restrictions in this model, compared to the separately estimated models for each data type, is only 3.93. Thus, we cannot reject identical indirect utility-difference parameters, providing the κ parameters are allowed to differ.

Table 3: Separate and Pooled Samples; Systematically Varying Utility Parameters

	1-ACT	2-MDC	3-OE	4-PC	5-MB	6-SP	Pooled
Utility parameters:							
exp(β_{11} +	7.516 (0.392)	0.293 (0.276)	-2.834 (-0.359)	1.496 (2.433)**	0.867 (2.165)**	1.545 (0.709)	1.588 (4.035)**
β_{12}*age$_i$	-3.012 (-0.519)	-0.322 (-2.448)**	-0.930 (-2.344)**	-0.398 (-3.752)**	-0.201 (-3.282)**	-0.580 (-1.427)	-0.449 (-6.340)**
β_{13}*purbin$_i$)*Opt C_i +	0.749 (0.114)	0.766 (0.754)	5.545 (0.704)	0.147 (0.348)	0.354 (1.096)	0.717 (0.421)	0.334 (1.148)
exp(β_{21})*Opt B_i +	-	-	-	-	-	0.198 (0.902)	0.713 (2.528)**
exp(β_{31})*Opt D_i +	-	-	-	-	-	-1.666 (-0.955)	-1.166 (-0.699)
exp(β_{41})*Opt E_i -	-	-	-	-	-	-16.431 (-0.067)	-15.306 (-0.095)
exp(β_{51} +	0.122 (0.091)	-2.154 (-4.734)**	-1.274 (-3.164)**	-0.738 (-2.191)**	-2.270 (-4.031)**	-2.911 (-2.715)**	-1.002 (-4.369)**
β_{52}*income$_i$	-28.038 (-2.028)**	-8.721 (-2.617)**	-5.475 (-2.858)**	-0.877 (-0.654)	1.265 (0.752)	-13.810 (-1.661)*	-4.582 (-4.310)**
β_{53}*age$_i$	-0.039 (-0.217)	0.056 (0.855)	-0.161 (-3.320)**	-0.155 (-3.338)**	-0.149 (-1.675)*	0.076 (0.827)	-0.073 (-2.947)**
β_{54}*male$_i$	-0.990 (-1.906)*	-0.085 (-0.530)	-0.140 (-0.955)	-0.032 (-0.285)	0.127 (0.781)	0.138 (0.336)	-0.071 (-1.092)
β_{55}*pownocc$_i$)*bid$_i$	-0.401 (-0.107)	2.830 (1.923)*	1.640 (1.296)	2.008 (2.055)**	2.200 (1.276)	5.271 (1.357)	2.078 (3.473)**
Ordered logit thresholds:							
α_0	-	-	-	-	-	-	-3.339 (-5.476)**
α_1	-	-	-	-	0.524 (6.836)**	-	-1.519 (-3.636)**
α_2	-	-	-	-	0.980 (9.722)**	-	0.068 (0.174)
α_3	-	-	-	-	1.845 (13.776)**	-	3.054 (4.209)**
κ multiples:[a]							
κ_1 (actual)	1.0	-	-	-	-	-	1.0
κ_2 (MDC)	-	1.0	-	-	-	-	1.620 (-3.256)**
κ_3 (OE)	-	-	1.0	-	-	-	2.448 (6.628)**
κ_4 (PC)	-	-	-	1.0	-	-	0.980 (-0.191)
κ_5 (MB)	-	-	-	-	1.0	-	3.507 (7.035)**
κ_6 ($SP_{pairwise}$)	-	-	-	-	-	1.0	1.720 (2.247)**
κ_7 (SP_{joint})	-	-	-	-	-	1.815 (2.081)**	3.072 (3.730)**
Fitted WTP at means of data							
Option C	n/a	$ 2.42	$ 0.54	$ 1.91	$ 5.84	$ 2.40	$ 1.90
Option B						6.82	6.15
Option D						1.06	0.94
Option E						0.00	0.00
H_0: var has no effect on WTP for Option C							
intercept	n/a	2.447 (2.271)**	-1.560 (-0.197)	2.234 (3.613)**	3.137 (5.183)**	4.456 (1.807)*	2.591 (6.759)**
income (in million $)		8.721 (2.617)**	5.475 (2.858)**	0.877 (0.654)	-1.265 (-0.752)	13.810 (1.661)*	4.582 (4.310)**
age (in decades)		-0.378 (-3.387)**	-0.768 (-1.953)*	-0.244 (-2.447)**	-0.052 (-0.682)	-0.656 (-1.606)	-0.376 (-5.660)**
male (dummy)		0.085 (0.530)	0.140 (0.955)	0.032 (0.285)	-0.127 (-0.781)	-0.138 (-0.336)	0.071 (1.092)
pownocc (% owner occ hous.)[c]		-2.830 (-1.923)*	-1.640 (-1.296)	-2.008 (-2.055)**	-2.200 (-1.276)	-5.271 (-1.357)	-2.078 (-3.473)**
purbin (% urbanized zip)		0.766 (0.754)	5.545 (0.704)	0.147 (0.348)	0.354 (1.096)	0.717 (0.421)	0.334 (1.148)

Derivatives of WTP for Option C at data means							
Intercept	n/a	$ -2.447	$ 1.560	$ 2.234	$ 3.137	$ 4.456	$ 2.591
income (in million $)		$ 0.375	$ 0.236	$ 0.038	$ -0.054	$ 0.594	$ 0.197
age (in decades)		$ -1.984	$ -4.033	$ -1.280	$ -0.275	$ -3.445	$ -1.974
male (dummy)		$ 0.046	$ 0.076	$ 0.017	$ -0.069	$ -0.075	$ 0.038
pownocc (% owner occ hous.)		$ -0.703	$ -0.408	$ -0.499	$ -0.547	$ -1.310	$ -0.517
purbin		$ 0.702	$ 5.081	$ 0.135	$ 0.324	$ 0.657	$ 0.306
Proportion WTP $6							
1 – ACT	0.290	-	-	-	-	-	0.20
2 – MDC	-	0.32	-	-	-	-	0.30
3 – OE	-	-	0.33	-	-	-	0.36
4 – PC	-	-	-	0.20	-	-	0.20
5 – MB	-	-	-	-	0.50	-	0.40
6 - $SP_{pairwise}$	-	-	-	-	-	0.34	0.31
7 - SP_{joint}	-	-	-	-	-	0.41	0.39
			-	-			
Number of choices employed	116	735	240	271	4840	1257	7459
Maximized Log L	-53.205	-433.490	-691.113	-671.361	-547.241	-243.866	-2670.611

[a] κ multiples are estimated as powers of e to ensure that they remain positive. For ease of interpretation, the point estimates have been exponentiated. The t-test statistics, however, remain tests of the hypotheses that the estimated exponents are zero. If the exponent is zero, the κ multiple is unity (exp(0)) and the dispersion for the sample in question is not different from that of the numeraire PDC sample.

[b] Percentages differ only because conditional dispersion of WTP distribution differs across data types. E[WTP] is constrained to be identical for all samples. Logistic distributional assumption is strong (especially in 3 - OE Tobit-type specification).

[c] The sum of the maximized log-likelihood values for the separately estimated models is –2640.277. There are 61 implicit parameters across the six different data types. There are only 21 parameters in this pooled-data model, implying a critical value for the likelihood ratio test of 55.759. The likelihood ratio test statistic for the restrictions embodied in this model is 60.67. This narrowly rejects the null hypothesis of identical preference parameters across these different elicitation methods.

Appendix

Non-response Analysis

In any survey-based research, it is important to consider the potential effects of unit non-response on the statistical relationships revealed from estimates based solely on respondents. Cameron, Shaw, and Ragland (1998) discusses the competing influences of survey topic salience and endogenous survey complexity in determining the probability that a given randomly selected mail survey recipient will elect to complete and return the survey instrument. Since the present study concerns passive use values, rather than active use values, it is spared most of the ills of endogenous complexity. In this case, the respondent's level of activity involving the goods to be valued does not affect the amount of time and/or energy required to fill out the questionnaire.

Fortunately, the type of elicitation method on any given individual's survey is completely unrelated to their individual attributes, since the elicitations are randomly assigned across participants. Thus the type of elicitation method cannot be correlated with any factor that might influence propensity to complete and return a questionnaire. Nevertheless, we retained the five-digit zip codes associated with every questionnaire mailed out. We matched each targeted household with the 1990 Census data for its zip code in an attempt to find any systematic differences in response propensities.

It is possible that respondents may be systematically less likely to cooperate with some types of surveys in our study than with others. If the observed nonresponse/response (a 0,1 variable) is regressed using a probit model on a set of dummy variables for elicitation mode (using the 2-MDC mode as the numeraire), we find no statistically significant differences in response rates across elicitation methods.

But we have explored the effects of Census zip code demographics on response propensities and have found that a wide variety of zip code level variables, used individually, make a statistically significant difference to response rates. PPUBINC (proportion on public assistance income) has a negative effect on response probability; IP200UP (proportion with income greater than 200% of the poverty level) has a positive effect on response probability; POWNOCC (proportion owner-occupied housing) has a positive effect on response probability, as does MEDGRRNT (median gross rent). PWHITE (proportion white) has a positive effect, and PBLACK (proportion black) has a negative effect. LOWED (proportion with low education levels) has a negative effect, and LANGIS (proportion language-isolated) also has a negative effect on response probabilities. Due to the high degree of collinearity among these variables, a model that includes all of them in the same specification does not reveal any individually statistically significant coefficients. The likelihood ratio test statistic for the hypothesis that all slope coefficients are jointly zero, however, is 115.8, which handily exceeds the critical value for a $\chi^2(8)$ distribution.

The relevant issue for comparisons of implied WTP values across elicitation modes, however, is whether there are systematic differences in response rates as a function of elicitation modes. Thus we have included not only dummy variables for elicitation modes and Census zip code demographics, but also a host of interaction terms between the two categories of variables. The relevant hypothesis test is a likelihood ratio test for this unrestricted model compared to a restricted model with no elicitation mode dummies or interaction terms. Results show that the maximized value of the probit log-likelihood is minimally compromised by restricting the elicitation-mode effects to zero, so we can safely argue that there are unlikely to be any quantitatively significant variations in response rates across elicitation modes.

Response rates do, however, differ according to demographic characteristics of the neighborhood (zip code) to which the survey was sent. Thus, in any model designed to estimate the overall social value of the GreenChoice[tm] program, it will be important to compensate for these differing response propensities. So far, we have implemented controls for differential response rates only crudely, by including as a the fitted response probability for each respondent as an additional explanatory variable. Presumably, simulating a 1.0 response probability would be one way to effect a crude correction for systematic variations in response rates across the sample.

Table A.1

Response Rates for Survey Variants

Survey Type	Phone 1-DC $6	Mail 2-DC $.50,1,2,4,6,9,12	Mail 3-OE	Mail 4-PC	Mail 5-MB	Mail 6-SP
Intended Sample	250	1400	500	500	900	600
Adjusted Sample (undeliv. no phone not NMPC customer)	199	1242	436	435	781	525
Completes	145	831	278	298	522	345
Response Rate (%)	70.4	66.9	63.8	68.5	66.8	65.7

Table A.2: More-general Utility Function; Non-Response Issue

	All data types	All data types	All data types	All data types
Type of Parameter	No regressors (review)	With regressors No Response Prob.	With Response Prob. No other regressors	With Response Prob. & other regressors
Utility parameters:[a]				
exp(β_{11} +	-0.310 (-1.40)	0.239 (0.672)	-1.081 (-1.412)	-0.177 (-0.550)
β_{12}*age$_i$	-	-0.442 (-6.466)**	-	-0.441 (-6.386)**
β_{13}*purbin$_i$	-	1.728 (3.813)**	-	1.828 (3.134)**
β_{14}*rprob$_i$) * Option C$_i$ +	-	-	1.459 (1.527)	0.119 (0.502)
exp(β_{21}	0.724 (2.32)**	0.737 (2.538)**	-2.075 (-0.581)	-1.603 (-0.437)
β_{22}*rprob$_i$) * Option B$_i$ +	-	-	4.248 (0.885)	2.915 (0.583)
exp(β_{31}	-1.397 (-0.64)	-1.269 (-0.665)	-20.638 (-0.805)	-10.021 (-0.499)
β_{32}*rprob$_i$) * Option D$_i$ +	-	-	25.857 (0.832)	11.943 (0.467)
exp(β_{41}	-19.86 (-0.07)	-9.812 (-0.117)	-5.692 (-0.107)	-7.569 (-0.067)
β_{42}*rprob$_i$) * Option E$_i$ -	-	-	-3.980 (-0.087)	-4.908 (-0.067)
exp(β_{51} +	-1.136 (-68.88)**	-1.835 (-6.255)**	-1.365 (-6.305)**	-2.036 (-5.334)**
β_{52}*income$_i$	-	-4.490 (-4.291)**	-	-4.441 (-4.113)**
β_{53}*age$_i$	-	-0.078 (-3.128)**	-	-0.075 (-2.967)**
β_{54}*male$_i$	-	-0.077 (-1.165)	-	-0.072 (-1.082)
β_{55}*pownocc$_i$	-	2.577 (4.053)**	-	2.861 (3.474)**
β_{56}*purbin$_i$	-	0.794 (4.311)**	-	0.799 (4.308)**
β_{57}*rprob$_i$) * bid$_i$	-	-	0.697 (2.192)**	-0.185 (-0.401)
Ordered logit thresholds:				
α_0	-3.285 (-4.61)**	-3.427 (-5.13)**	-4.256 (-3.55)**	-2.701 (-2.20)**
α_1	-1.481 (-3.37)**	-1.534 (-3.34)**	-1.925 (-2.93)**	-1.213 (-1.98)**
α_2	0.09364 (0.25)	0.122 (0.27)	0.108 (0.25)	0.0954 (0.29)
α_3	3.047 (3.72)**	3.248 (3.93)**	3.922 (3.19)**	2.564 (2.08)**
κ multiples:[b]				
κ_1 (ACT)	1.0	1.0	1.0	1.0
κ_2 (MDC) (numeraire)	1.732 (2.85)**	1.621 (3.03)**	2.223 (3.01)**	1.278 (0.55)
κ_3 (OE)	2.575 (5.22)**	2.316 (5.66)**	3.295 (4.71)**	1.834 (1.41)
κ_4 (PC)	0.9505 (-0.30)	0.972 (-0.22)	1.245 (0.89)	0.7676 (-0.61)
κ_5 (MB)	3.514 (5.83)**	3.695 (6.81)**	4.535 (5.45)**	2.918 (2.36)**
κ_{6P} (SP$_P$)	1.735 (2.01)**	1.774 (2.28)**	2.246 (2.52)**	1.359 (-0.63)
κ_{7J} (SP$_J$)	3.236 (3.54)**	3.136 (3.71)**	4.304 (3.79)**	2.458 (1.69)
Fitted WTP at means of data				
Option C	$ 2.28	$ 1.85	$ 2.20	$1.82
Option B	$ 6.42	$ 6.35	$ 41.35	$26.46
Option D	$ 0.77	$ 0.85	$ really huge	$10590.11
Option E	$ 0.00	$ 0.00	$ 0.18	$0.15
H_0: var has no effect on WTP for Option C				
income (in million $)	-	4.490 (4.291)**	-	4.441 (4.113)**
age (in decades)	-	-0.364 (-5.692)**	-	-0.366 (-5.643)**
male (dummy)	-	0.077 (1.165)	-	0.072 (1.082)
pownocc (% owner occ hous.)[c]	-	-2.577 (-4.053)**	-	-2.861 (-3.474)**
purbin (% in urban areas)	-	0.933 (2.075)**	-	1.029 (1.788)*
rprob (resp prob prior model)	-	-	0.761 (0.918)	0.304 (0.739)
Derivatives of WTP for Option C at data means				
income (in million $)	-	$ 0.193	-	$ 0.191
age (in decades)	-	-1.908	-	-1.920
male (dummy)	-	0.042	-	0.039
pownocc (% owner occ hous.)	-	-0.641	-	-0.711
purbin (% in urban areas)	-	0.855	-	0.943
rprob (resp prob prior model)	-	-	$ 0.505	0.202

Proportion WTP $6 [d]				
1 - ACT	0.233	0.204	0.176	0.251
2 - MDC	0.334	0.301	0.333	0.298
3 - OE	0.386	0.357	0.385	0.355
4 - PC	0.222	0.197	0.225	0.194
5 - MB	0.233	0.409	0.416	0.407
6_P - SP_P	0.335	0.317	0.335	0.309
6_J - SP_J	0.409	0.393	0.411	0.391
Number of choices employed	7459	7459	7459	7459
Maximized Log _	-2714.733	-2659.618	-2712.259	-2659.022 [f]

[a] The midpoint of the respondent's reported income bracket is included as a proxy for "socioeconomic status" which is assumed to influence preferences. Income as a determinant of the budget constraint drops out of the indirect utility-difference function in the linear forms employed in this study.

[b] κ multiples are estimated as powers of e to ensure that they remain positive. For ease of interpretation, the point estimates have been exponentiated. The t-test statistics, however, remain tests of the hypotheses that the estimated exponents are zero. If the exponent is zero, the κ multiple is unity (exp(0)) and the dispersion for the sample in question is not different from that of the numeraire PDC sample.

[c] Variables **pownocc** and **purbin** are some of the 1990 Census data merged with the samples by 5-digit zip codes. These are interpreted as neighborhood characteristics. The **rprob** variable is a fitted response probability from a preliminary probit model that uses a variety of 1990 Census variables at the 5-digit zip code level, interacted with survey type, to explain the response/nonresponse decisions of all households who were mailed a questionnaire.

[d] Proportions differ only because conditional dispersion of WTP distribution differs across data types. E[WTP] is constrained to be identical for all samples. Logistic distributional assumption is strong (especially in 3 - OE Tobit-type specification).

[e] None of the **rprob** coefficients is individually statistically significant. A likelihood ratio test for the null hypothesis that all **rprob** coefficients are jointly zero is distributed $\chi^2(5)$ with a 5% critical value of 11.07. The test statistic does not achieve that value, so the hypothesis that response probabilities have no effect on E[WTP] cannot be rejected.

[f] The five **rprob** coefficients in this model are not jointly statistically different from zero.