WP 2001-18
September, 2001

Working Paper

Department of Agricultural, Resource, and Managerial Economics
Cornell University, Ithaca, New York 14853-7801 USA

Does Specification Error Explain the Discrepancy Between Open-Ended and Dichotomous Choice Contingent Valuation Responses? A Comment on "Monte Carlo Benchmarks for Discrete Valuation Methods" by Ju-Chin Huang and V. Kerry Smith

By Gregory L. Poe and Christian A. Vossler

Does Specification Error Explain the Discrepancy Between Open-Ended and Dichotomous Choice Contingent Valuation Responses? A Comment on "Monte Carlo Benchmarks for Discrete Valuation Methods" by Ju-Chin Huang and V. Kerry Smith.

Gregory L. Poe[a]

454 Warren Hall
Department of Applied Economics and Management
Cornell University
Ithaca, NY 14853-7801
(607) 255-4707
Email: glp2@cornell.edu

and

Christian A.Vossler

157 Warren Hall
Department of Applied Economics and Management
Cornell University
Ithaca, NY 14853-7801
(607) 255-2085
Email: cav22@cornell.edu

September 12, 2001

Abstract:

In this working paper we demonstrate that some of the statistical tests used by Huang and Smith in a recent *Land Economics* article (74(2 1998): 186-202) were erroneous, and raise concerns about their corresponding conclusions. Specifically, using data from one of the studies that they showcase, we demonstrate that Huang and Smith's analysis suggesting statistical equality between hypothetical dichotomous choice responses and actual contributions is incorrect. We further show that their purported equality between dichotomous choice and open-ended response formats is unfounded. Based on these analyses we conclude that when real humans make real or stated decisions, the observed procedural variance across elicitation methods and the degree of hypothetical bias are more fundamental than relying on alternative econometric specifications.

[a] Poe and Vossler are Associate Professor and Graduate Research Assistant, respectively. The authors are listed alphabetically – senior authorship is shared. We wish to acknowledge the helpful comments of Ian Langford and Bill Schulze on this research, although responsibility for errors remains our own. This comment was written, in part, while Poe was a Visiting Fellow at the Jackson Environmental Institute (JEI) and the Centre for Social and Economic Research on the Global Environment (CSERGE) at the University of East Anglia. Funding for this research was provided by USDA Regional Project W-133 and Cornell University. Cornell University AEM Working Paper 2001-18.

I. INTRODUCTION

In a recent article, Huang and Smith (1998, hereafter HS) use Monte Carlo simulation methods to suggest that the procedural variance observed between open-ended (OE) and dichotomous choice (DC) contingent valuation (CV) responses can be attributed to specification error in modeling the DC responses. In particular, HS argue that employing alternative specifications of the error term for DC responses can provide estimates of mean willingness to pay (WTP) that are "virtually identical to the mean of the raw data derived with open-ended CV and not significantly different from the mean…for actual purchases" (p. 191). Hence, they claim that the evidence from a large body of laboratory and field research that DC-CV question formats yield substantially larger estimates of mean WTP than OE response methods is "unfounded" (p. 187). While we applaud the Monte Carlo methods used by HS to demonstrate that error specification is important in providing unbiased and accurate estimates of WTP and conditional WTP, we wish to caution the reader that alternative specifications of the error term are not likely to bridge the gap between OE and DC WTP estimates. When real humans make real or stated decisions, observed procedural variance across elicitation methods and the degree of hypothetical bias are more fundamental than relying on alternative econometric specifications.

To demonstrate this point we roughly follow the organization of the HS paper. In the following section we provide a brief review of the Balistreri et al. (2001) data showcased by HS[i]. This data is used to demonstrate that, in contrast to the HS paper, the mean WTP estimate from DC-CV data is significantly different from actual contributions and that the DC and OE distributions are significantly different from each other. In the third section we raise concerns about the functional forms, error distributions, and welfare estimates used in the Monte Carlo analysis of HS. Using a broader range of utility-theoretic specifications than the linear logistic and probit models employed by HS, we demonstrate that employing alternative error specifications is not likely to overcome the observed disparity in mean WTP values across elicitation methods. The fourth section addresses concerns about the Turnbull lower bound estimator used by HS in support of their not significantly different and virtually identical claims, and the increased application of this method to provide “conservative” estimates of mean WTP from DC-CV responses. We conclude with some final thoughts on relying on alternative error specifications, rather than seeking a better understanding of human generated responses, to measure and correct for procedural variances observed in applied economic research.

II. ON “NOT SIGNIFICANTLY DIFFERENT” AND “VIRTUALLY IDENTICAL”

The Balistreri et al. study compared mean WTP values obtained from an English Auction (using real money) with hypothetical DC and OE survey responses for an insurance policy against a known loss (\$10) with a known probability (40%). Participants were endowed with \$80, being actual or hypothetical money depending on whether actual or hypothetical WTP values are elicited. In Table 2 of their paper, HS provide a partial summary of the Balistreri et al. results. Actual values elicited from an English Auction provide a mean WTP of \$3.66 with a standard deviation of 1.15 (n=52), which is slightly, but significantly, below the expected value of \$4 for such an insurance policy (t=2.13, p=0.04)[ii]. In such an auction prices are raised sequentially until the bidding stops (i.e., only one active bidder remains). This method is relatively transparent and incentive compatible for private goods (Davis and Holt, 1993) and hence serves as a reference point for assessing hypothetical bias. In this same table, the mean derived from the hypothetical OE responses is \$4.58 (standard deviation = \$5.38, n=345). Mean WTP values derived from DC responses, wherein the price of the insurance policy is varied across respondents is \$5.77 (standard error = \$0.26) for the non-negative mean of a linear logistic WTP function and \$5.63 (standard error = \$0.26) using the entire linear logistic distribution, including possible negative values[iii]. Using methods described in Haab and McConnell (1997) HS further estimate the non-parametric Turnbull lower bound estimate of the mean (\$4.56, standard error =\$0.31). Note that in presenting these results we are careful to distinguish between the standard deviation of the distribution and the standard error of the mean.

The above statistics provide enough information to assess the validity of the HS claim that DC responses provide a mean WTP estimate that is not significantly different from the estimated mean for actual purchases. Unfortunately, in making this claim, HS do not provide information about how this conclusion was reached. One may conjecture from the statistics provided in Table 2 of HS that a difference of means t-test was applied. A widely adopted form of this test, which accommodates unequal sample variances from two independent samples, is known as 'Welsh's approximate t'. The test statistic is:

$$\mathrm{t}_{v} = \frac{\eta_1 - \eta_2}{\sqrt{\mathrm{SE}_1^2 + \mathrm{SE}_2^2}},$$

where η_1 and η_2 denote the parameter estimates, and SE_1 and SE_2 denote the estimated standard errors (Zar, 1996). The test statistic follows a Student's t-distribution with degrees of freedom (d.f.) approximated by:

$$v = \frac{(\mathrm{SE}_1^2 + \mathrm{SE}_2^2)^2}{\frac{(\mathrm{SE}_1^2)^2}{n_1 - 1} + \frac{(\mathrm{SE}_2^2)^2}{n_2 - 1}}.$$

When this difference of means test is applied to the summary data from the Balistreri et al. study, we reach exactly the opposite conclusion than that reached by HS – *even when the lowest* *possible estimate of the DC response function is used (i.e., the Turnbull lower bound estimate), the mean WTP estimate is significantly higher than actual contributions.* The t statistic from this test is 2.58 (d.f. = 425.36) resulting in a significance level of p = 0.01. As such, we maintain that the "not statistically different" claim made by HS is itself unfounded.

In contrast, we concur with HS that the mean WTP estimate derived from DC data using the Turnbull Lower Bound method is virtually identical to the raw mean obtained from the OE data. Indeed, if anything, the point estimate for the Turnbull Lower Bound estimate is lower than that for the OE data. But does this measure of central tendency really reflect the underlying differences in the distributions? We think not.

In addressing this issue, it is helpful to have additional information on the distributions of responses from these two formats. Table 1 replicates Table 3 in Balistreri et al. (2001): the first column indicates the posted price or bid values used in the DC questionnaire, the second column indicates the number of responses obtained at each bid value, and the third column provides the proportion of DC respondents that "accepted" the posted price. The last two columns report the estimated, rounded number of OE respondents, and corresponding proportions, that would have answered yes to each DC value, assuming that respondents would have chosen to buy the insurance if the posted price had been less than or equal to their OE values[iv]. These proportions, along with the survival, or reverse cumulative, distribution, of the OE responses are depicted graphically in Figure 1. For reference purposes all OE responses are provided in Appendix 2.

From Table 1 and Figure 1, it should be readily apparent that the DC responses stochastically dominate the OE responses. At some of the posted prices the probability difference between the two methods is relatively small (e.g., 5.62% at $1), while at others it is quite substantial (e.g, 16.14% at $6). Regardless of magnitude, the fact remains that at every point at which we have a possible comparison, the OE survival function lies below the DC survival function.

Stochastic dominance need not imply significance. Unfortunately, we have a problem of comparability. The OE survival function is continuous while the DC survival distribution is not. To make these two sets of data comparable, either both have to be converted to continuous distributions or both have to be converted to discrete functions. Balistreri et al. followed the former approach, rejecting the null hypothesis of equality between OE and DC ($p<0.01$, $\chi^2(2)=15.11$) using linear logistic error specifications. Here, we make use of the converted OE values provided in Table 1 and conduct a Kolmogorov-Smirnov type test. We return to comparisons associated with continuous distributions in the following section.

The Smirnov Test (Conover, 1980) can be used to test whether two empirical distributions are equal, when the distributions are derived from two mutually independent, random samples. Because of the discrete nature of our data, however, this test is conservative (Noether, 1967). The Smirnov Test statistic is exactly the same as the Kolmogorov-Smirnov *D*-statistic:

$$D = \max | F(x) - G(x) |$$

where F(x) and G(x) depicts the DC and OE distributions, respectively. The maximum distance between distributions is 0.1614 and occurs at \$6. Applying the appropriate formula in Conover (1980, p. 473), the large-sample approximation for the critical $D_{0.01}$ value is 0.12 and so we reject the hypothesis of equal distributions beyond the 1% significance level.

Given these statistical results and the observation that the DC distribution stochastically dominates the OE response function, why then did HS get a result in which the estimated mean WTP for the DC responses lies below, but not significantly so, the OE mean WTP? In part the answer lies in the fact that the HS article only considers alternative specifications for the DC responses and took the raw mean from the OE as given.

However, fair and complete tests of comparability necessitates that we consider shortcomings and modifications to both sides of the comparison. Recall, that the raw mean of the data is drawn from a simple average of WTP values reported to avoid the risk of a 40% chance of losing 10 dollars. Inspection of Figure 1 and Appendix 2 indicate that 16 (or 4.63%) of the OE observations exceeded the highest possible loss of \$10[v] with six observations at \$30 or higher. These extreme, "irrational" values exert a strong influence on the mean and the variance of the OE responses, both of which are critical to the standard difference of means test. Following experimental economics standards that all values be retained, regardless of whether they appear irrational or not, Balistreri et al. used the entire data set in calculating the mean WTP of \$4.58. In making this decision, they note however, that the irrationally high "bids, as is typically done in CV studies might justifiably be trimmed" (p. 281).

For demonstrative purposes, rather than dropping these observations entirely from the data set, we recoded these elevated values to the highest "rational" response of \$10. In this case the mean WTP falls to \$3.88 (s.d = 2.61, n = 345). This estimated mean is not significantly different from the expected value of \$4 (t=0.84, p.0.40) nor is it different from the English Auction results (t=1.05, p.0.30). It is however, marginally different than the Turnbull estimate (t=1.99, p.0.05). Alternatively, in an effort to ensure comparability between elicitation formats, these extreme values might be recoded to \$12, the highest value in the DC bid vector. Under these conditions the estimated mean is \$3.97, which is not significantly different from the expected value of \$4 (t=0.17, p.0.87) nor the value obtained from the English Auction (t=1.42, p.0.16). This value is still marginally different than the Turnbull lower bound estimate (t=1.69, p.0.09), lying between the 5 and 10 percent level of significance.

III. ON FUNCTIONAL FORM, ERROR DISTRIBUTIONS, AND WELFARE ESTIMATES

Omission of relevant explanatory variables and misspecification of the functional relationship between the dependent variable and explanatory variables are common econometric problems that can lead to erroneous economic conclusions. The Monte Carlo results of HS illustrate this well-known finding in the specific context of DC-CV. However, the algebraic model HS use to benchmark the degree of specification error is suspect.

Our caution about applying the HS results stems from the fact that the algebraic model used by HS to demonstrate specification error is inconsistent with the structure of their Monte Carlo experiment. By construction, their various preference specifications restrict the utility difference to be positive. However, in some 924 of their 4800 cases “technically feasible but economically implausible…negative use values” did occur (p. 197). As described in their footnote 18, such observations were deleted. The potential problem with such a selected simulation approach arises because the linear logistic and probit models used in the subsequent HS analysis are unbounded, including possible negative values. As pointed out by Haab and McConnell (1998) in the same issue of this journal, “if the distribution of WTP is known to have lower and upper bounds which are narrower than implied from the estimation, then the initial model is misspecified and the parallel estimates are inefficient, failing to use all the available information and inconsistent from assuming the wrong distribution of WTP” (p. 217). As such, the HS demonstration of specification errors may itself be associated with the fact that they chose specifications that are not consistent with their underlying experiment[vi].

Our concern here is whether a similar misspecification in the Balistreri et al. paper, wherein an unbounded linear logistic function was applied to data that was unambiguously non-negative, could have led to erroneous rejections of equality between estimated mean WTP values from the DC and OE data. To investigate this possibility we reestimate the DC-WTP relationship using specifications of the error term that are consistent with utility-theoretic restriction that the utility difference be non-negative (see Hanemann and Kanninen, 1999). In arriving at these estimates, the structure of the laboratory experiments precluded using many different specifications for the algebraic model[vii]. Hence, we restrict ourselves to an algebraic model that specifies the yes/no response choice as a function of a constant term and either the bid or the natural log of the bid. Using logistic, normal and two-parameter Weibull error distributions, we further impose a theoretically desirable constraint on the upper bound of the estimated WTP distribution. Each experiment participant is (hypothetically) endowed with $80 and so estimated WTP should fall at or below $80. However, based on OE responses, the upper bound of estimated WTP is likely to be closer to $40. We employ two approaches to imposing this upper bound restriction. First, we normalize/truncate the estimated cdf using the approach of Boyle, Welsh, and Bishop (1988). Second, we impose the restriction that an individual's WTP lie between zero and $40 directly into the econometric model through a technique referred to as "pinching" (Ready and Hu, 1995). Finally, we abandon all algebraic model and error distribution assumptions and use Kriström's (1990) nonparametric approach and the Turnbull lower bound estimate (see Haab and McConnell, 1997). Using linear interpolation, the upper bound for the Kriström nonparametric estimator is $15.60. The parameter estimates for the various specifications we explore are included as Appendix 2.

Overall, we obtain eleven different mean WTP estimates and report these values - along with the linear logistic and probit estimates - in Table 2. Standard errors and 95% confidence intervals for the parametric specifications are estimated using the Krinsky and Robb procedure with 10,000 random draws (see Park, Loomis, and Creel, 1991). Standard errors and confidence intervals for the non-parametric specifications are calculated using formulas provided in the literature (Haab and McConnell, 1997). Empirical distributions of mean WTP for the OE, English Auction, and non-parametric specifications were generated from the respective sample means and standard errors. The convolutions method (Poe, Severance-Lossin, and Welsh, 1994) is used to conduct statistical tests under the null hypothesis that the mean WTP estimates from the raw OE data are equal to corresponding estimates obtained from the DC responses.

As indicated in Table 2, the hypothesis of identical OE and DC mean WTP can be rejected for all specifications at the 5% significance level or greater, with the sole exception being the Turnbull lower bound estimate. Not surprisingly, all DC mean WTP estimates are statistically different than the English Auction estimates. Note, in particular, that this bootstrapping of means approach corroborates the earlier parametric comparisons of Turnbull and English auction estimates.

IV. ON LOWER BOUND APPROACHES

To this point we have merely used the statistics provided by HS and a reexamination of the Balistrei et al. data to refute the “statistical different” and “virtually identical” statements made by HS and to express our concerns about the Monte Carlo simulations. Under a broad range of specifications we found that the HS claims cannot be supported. The only instance in which equality does appear to hold across elicitation methods is when the most extreme lower bound assumption regarding DC responses is made. Here we raise particular concerns about this estimator and its increased use in CV.

The application of the Turnbull lower bound approach in CV appears to have arisen out of “highly publicized damage assessment cases” (Haab and McConnell, 1997) and the corresponding desire to have a legalistically defendable, conservative estimate of hypothetical WTP (see Harrison and Kristr m̄, 1995). Briefly, this estimator masses all the positive WTP responses at the corresponding DC value (for a more detailed presentation see Haab and McConnell, 1997), rather than assuming that the distribution of WTP includes values that lie between DC levels.

Our hesitation towards the increased application of this method arises out a number of interrelated concerns. First, while we agree that the Turnbull estimate is relatively transparent, uses only the information provided and could, perhaps, be regarded as the “minimum legal” WTP from implicit DC “contracts” between the researcher and the respondent (Harrison and Kriström, 1995), we maintain that the goal of CV should be to provide the best, rather than lower bound, estimate of WTP. When hypothetical bias is found to exist, we argue that there is a greater need to explore how and why respondents provide answers that appear “inconsistent” with actual contributions instead of relying on technical, and as we demonstrate below somewhat arbitrary, econometric permutations to bring hypothetical DC values down to apparently reasonable levels. That is, our efforts should be directed towards developing question formats that help respondents provide more realistic representations of their underlying WTP. Some recent, promising modifications to the DC methods along these lines are presented by Champ et al. (1997), Poe and Welsh (1998), Cummings and Taylor (1999) and Ready, Navrud, and Dubourg (2001).

We also question the apparent equating of the terms "distribution free" and "assumption free" that occurs by some defenders of the Turnbull approach. By massing points at the DC values rather than, say, assuming the values to be distributed between DC bids as was done by Kriström (1990), the modeler is making the rather strong assumption that all values can be massed at their corresponding DC bid function. Examination of the OE WTP distribution in Figure 1 shows that such an assumption is counterfactual. That making such an assumption grossly influences estimated mean values is demonstrated by applying the Turnbull lower bound estimator to the converted OE responses in Table 1. Under these assumptions[viii] the mean WTP value is estimated to be \$3.25 (standard error = \$0.25) which is lower than the mean WTP for the English Auction, but not significantly so (t=1.39, p.0.17). However, in stark contrast to HS, imposing this parallel assumption on the OE data engenders a highly significant difference between the Turnbull lower bound DC estimate and OE mean WTP (t=3.30, p<0.01).

Additional concern about using the Turnbull estimator as providing a lower bound estimate of WTP is that it is highly dependent upon the bid vector, a point raised in Haab and McConnell (1997) and empirically demonstrated here. In turn this dependence "suggests caution with respect to absolute interpretations of the welfare measure" to be the lower bound estimate (Haab and McConnell, 1997, p. 259). To demonstrate this point, we start with the full bid design used in Balistreri et al.[ix], explore the effects on mean WTP associated with dropping one of the bid levels (and the corresponding responses) from the data set, and compare the resulting values with those obtained from the Kriström (1990) specification, which masses values equally across the bid interval, and a series of non-negative "pinched" parametric distributions. The results from this exercise are provided in Table 3. As shown, relative to the full bid vector the "jackknifed" bid vectors lower the Turnbull lower bound estimates for each alternative, sometimes substantially. In contrast, the corresponding measures of WTP derived from the Kriström and parametric approaches tend to vary around the full bid vector value and exhibit a lot less fluctuation. Using the full bid vector as a reference point, the jackknifed Turnbull Lower Bound estimates exhibit a much higher mean squared error (0.30) on average than that of the Kriström (0.04) and the continuous parametric distributions (0.04 to 0.07). As such, the near-perfect alignment of the mean OE and the corresponding Turnbull estimate from DC responses appears to be a serendipitous result particular to the bid design used in Balistreri et al.

In summary, we have substantial concerns about the estimator that HS used to support their not statistically significant and virtually identical claims, and broader concerns about the increased use of this estimator in CV research. Adopting equal, counterfactual assumptions for the OE responses drives a wedge between the mean OE and DC estimates. Further, Turnbull lower bound estimates are extremely dependent upon the bid vector, to the extent that they may be regarded as somewhat arbitrary values. It appears that estimation methods that assume a continuity in values are less susceptible to changes in the bid structure.

V. SUMMARY AND CONCLUSIONS

In the abstract of their paper, HS maintain that the “belief that discrete contingent valuation questions yield substantially larger estimates of the mean (and median) willingness to pay (WTP) for nonmarket resources is unfounded” (p. 186). This claim is purportedly supported by their reassessment of the results from specific studies on elicitation effects. Drawing WTP values from known distributions, they then conduct Monte Carlo simulations to show that the degree of error associated with commonly used DC response functions can “easily span the differences between” OE and DC results (p. 200).

Using data from one study showcased by HS, we show that reasonable, and correct, statistical comparisons refute their statements that respecifications can provide DC values that are virtually identical to OE responses and not statistically different from actual WTP. While we applaud their efforts to demonstrate the importance of specification error and omitted variable bias in the estimating WTP, our close examination of the one data set that they use to support their claims and that is available to us, leads us to conclude that assuming simulated individuals and employing creative econometrics may provide some useful insights on the expected magnitude of the difference, but will not obviate the fundamental observation that a disparity occurs between DC and OE mean WTP values. Human subjects reporting real and hypothetical values apparently demonstrate behavioral tendencies that lead to hypothetical bias and procedural variance. Rather than assuming away these behaviors, a more promising research agenda would be to increase our understanding as to why these systematic differences occur and to develop elicitation methods that account for these sorts of variation.

Appendix 1. Distribution of Open Ended Responses

Obs	Value	Obs	Value	Obs	Value	Obs	Value	Obs	Value	Obs	Value	Obs	Value	Obs	Value	Obs	Value
1	0	41	0.5	81	2	121	3	161	4	201	4	241	5	281	5	321	10
2	0	42	0.5	82	2	122	3	162	4	202	4	242	5	282	5	322	10
3	0	43	0.5	83	2	123	3	163	4	203	4	243	5	283	5	323	10
4	0	44	1	84	2	124	3	164	4	204	4	244	5	284	5	324	10
5	0	45	1	85	2	125	3	165	4	205	4	245	5	285	5	325	10
6	0	46	1	86	2	126	3	166	4	206	4	246	5	286	5	326	10
7	0	47	1	87	2	127	3	167	4	207	4	247	5	287	5	327	10
8	0	48	1	88	2	128	3	168	4	208	4	248	5	288	5	328	10
9	0	49	1	89	2	129	3	169	4	209	4	249	5	289	5	329	10
10	0	50	1	90	2	130	3.1	170	4	210	4	250	5	290	5.05	330	12
11	0	51	1	91	2	131	3.2	171	4	211	4	251	5	291	5.25	331	15
12	0	52	1	92	2	132	3.25	172	4	212	4	252	5	292	5.25	332	15
13	0	53	1	93	2	133	3.5	173	4	213	4	253	5	293	5.5	333	20
14	0	54	1	94	2	134	3.5	174	4	214	4	254	5	294	6	334	20
15	0	55	1	95	2	135	3.5	175	4	215	4	255	5	295	6	335	20
16	0	56	1	96	2	136	3.5	176	4	216	4	256	5	296	6	336	20
17	0	57	1	97	2	137	3.5	177	4	217	4	257	5	297	6	337	20
18	0	58	1	98	2	138	3.5	178	4	218	4	258	5	298	6	338	26.6
19	0	59	1	99	2	139	3.5	179	4	219	4.2	259	5	299	6	339	26.76
20	0	60	1	100	2	140	3.99	180	4	220	4.25	260	5	300	6	340	30
21	0	61	1	101	2	141	3.99	181	4	221	4.5	261	5	301	6.5	341	30.2
22	0	62	1	102	2.01	142	3.99	182	4	222	4.5	262	5	302	6.66	342	32
23	0	63	1.25	103	2.5	143	3.99	183	4	223	4.5	263	5	303	6.75	343	35
24	0	64	1.5	104	2.5	144	3.99	184	4	224	4.5	264	5	304	6.78	344	40
25	0	65	1.5	105	2.5	145	4	185	4	225	4.5	265	5	305	7	345	40
26	0	66	1.5	106	2.5	146	4	186	4	226	4.5	266	5	306	7		
27	0	67	1.5	107	2.5	147	4	187	4	227	5	267	5	307	7		
28	0	68	1.5	108	2.5	148	4	188	4	228	5	268	5	308	7		
29	0	69	1.5	109	2.5	149	4	189	4	229	5	269	5	309	7		
30	0	70	1.5	110	2.5	150	4	190	4	230	5	270	5	310	7.49		
31	0	71	2	111	2.5	151	4	191	4	231	5	271	5	311	7.5		
32	0	72	2	112	2.5	152	4	192	4	232	5	272	5	312	7.5		
33	0.01	73	2	113	2.5	153	4	193	4	233	5	273	5	313	7.5		
34	0.01	74	2	114	2.5	154	4	194	4	234	5	274	5	314	7.5		
35	0.01	75	2	115	2.5	155	4	195	4	235	5	275	5	315	8		
36	0.1	76	2	116	3	156	4	196	4	236	5	276	5	316	8		
37	0.2	77	2	117	3	157	4	197	4	237	5	277	5	317	9		
38	0.4	78	2	118	3	158	4	198	4	238	5	278	5	318	9		
39	0.5	79	2	119	3	159	4	199	4	239	5	279	5	319	9.99		
40	0.5	80	2	120	3	160	4	200	4	240	5	280	5	320	9.99		

Appendix 2. Parameter Estimates for Various Willingness to Pay Functions

Distribution	Constant (s.e.)	Slope (s.e.)	χ^2	Log-L	Pseudo R^2
Logistic	2.6876 (0.2744)	-0.4771 (0.0501)	144.97	-214.90	0.2522
Log-Logistic	3.4234 (0.4082)	-2.1094 (0.2441)	144.83	-214.96	0.2520
Pinched Logistic	4.4199 (0.7273)	-2.4962 (0.4178)	148.04	-213.36	0.2576
Weibull	2.9802 (0.2906)	-1.5641 (0.1579)	147.76	-213.50	0.2571
Pinched Weibull	3.5017 (0.4139)	-1.7269 (0.2182)	147.94	-213.41	0.2574
Normal	1.5790 0.1486	-0.2763 (0.0263)	143.98	-215.39	0.2505
Log-Normal	1.8604 (0.1941)	-1.1575 (0.1172)	141.77	-216.49	0.2467
Pinched Normal	2.7454 (0.5268)	-1.5472 (0.2985)	146.34	-214.21	0.2546

Appendix 3. Likelihood Functions and Coefficient Estimates for Various Parametric Forms Used in This Comment. (Note: U_i denotes the upper bound; Y_i=1 if $WTP_i \geq bid_i$):

Logistic

$$\ln L = \sum_{i=1}^{n} \left\{ Y_i \ln\left(\frac{1}{1+e^{-\alpha-\beta * bid_i}} \right) + (1-Y_i)\ln\left(1 - \frac{1}{1+e^{-\alpha-\beta * bid_i}} \right) \right\}$$

Log-Logistic

$$\ln L = \sum_{i=1}^{n} \left\{ Y_i \ln\left(\frac{1}{1+e^{-\alpha-\beta * \ln(bid_i)}} \right) + (1-Y_i)\ln\left(1 - \frac{1}{1+e^{-\alpha-\beta * \ln(bid_i)}} \right) \right\}$$

Pinched Log-Logistic

$$\ln L = \sum_{i=1}^{n} \left\{ Y_i \ln\left[\left(\frac{1}{1+e^{-\alpha-\beta * \ln(bid_i)}} \right)\left(1 - \frac{bid_i}{U_i} \right) \right] + (1-Y_i)\ln\left[1 - \left(\frac{1}{1+e^{-\alpha-\beta * \ln(bid_i)}} \right)\left(1 - \frac{bid_i}{U_i} \right) \right] \right\}$$

Weibull

$$\ln L = \sum_{i=1}^{n} \left\{ Y_i \ln\left(e^{-e^{-\alpha-\beta * \ln(bid_i)}} \right) + (1-Y_i)\ln\left(1 - e^{-e^{-\alpha-\beta * \ln(bid_i)}} \right) \right\}$$

Pinched Weibull

$$\ln L = \sum_{i=1}^{n} \left\{ Y_i \ln\left[\left(e^{-e^{-\alpha-\beta * \ln(bid_i)}} \right)\left(1 - \frac{bid_i}{U_i} \right) \right] + (1-Y_i)\ln\left[1 - \left(e^{-e^{-\alpha-\beta * \ln(bid_i)}} \right)\left(1 - \frac{bid_i}{U_i} \right) \right] \right\}$$

Normal

$$\ln L = \sum_{i=1}^{n} \left\{ Y_i \ln[\Phi(\alpha + \beta * bid_i)] + (1-Y_i)\ln[1 - \Phi(\alpha + \beta * bid_i)] \right\}$$

Log-Normal

$$\ln L = \sum_{i=1}^{n} \left\{ Y_i \ln[\Phi(\alpha + \beta * \ln(bid_i))] + (1-Y_i)\ln[1 - \Phi(\alpha + \beta * \ln(bid_i))] \right\}$$

Pinched Log-Normal

$$\ln L = \sum_{i=1}^{n} \left\{ Y_i \ln\left[\Phi(\alpha + \beta * \ln(bid_i))\left(1 - \frac{bid_i}{U_i} \right) \right] + (1-Y_i)\ln\left[1 - \Phi(\alpha + \beta * \ln(bid_i))\left(1 - \frac{bid_i}{U_i} \right) \right] \right\}$$

References

Balistreri, E., G. McClelland, G. L. Poe, and W. D. Schulze, 2001. "Can Hypothetical Questions Reveal True Values? A Laboratory Comparison of Dichotomous Choice and Open-Ended Contingent Values with Auction Values", *Environmental and Resource Economics,* 18: 275-292.

Boyle, K. J., M. P. Welsh, and R. C. Bishop, "Validation of Empirical Measures of Welfare Change: Comment", *Land Economics*, 64(1 Feb): 94-98.

Champ, P. A., R. C. Bishop, T. C. Brown, and D. W. McCollum, 1997. "Using Donation Mechanisms to Value Nonuse Benefits from Public Goods", *Journal of Environmental Economics and Management*, 33(2 June): 151-163.

Conover, W. J., 1980. **Practical Nonparametric Statistics**. 2d ed. New York: John Wiley & Sons.

Cummings, R. G., and L. O. Taylor, 1999. "Unbiased Value Estimates of Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method", *American Economic Review*, 89(3 June): 649-665.

Davis, D. D. , and C. A. Holt, 1993. **Experimental Economics**, New Jersey, Princeton University Press.

Haab, T. C, and K. E. McConnell, 1997. Referendum Models and Negative Willingness to Pay: Alternative Solutions". *Journal of Environmental Economics and Management*, 32 (2 Feb):251-70.

Haab, T. C., and K. E. McConnell, 1998. "Referendum Models and Economic Values: Theoretic, Intuitive, and Practical Bounds on Willingness to Pay", *Land Economics*, 74(2 May):186-202.

Hanemann, M., and B. Kanninen, 1999. "Chapter 11: The Statistical Analysis of Discrete-Response CV Data", in I. J. Bateman and K. G. Willis eds., **Valuing Environmental Preferences**, Oxford University Press, Oxford.

Harrison, G. W., and B. Kriström, 1995. "On the Interpretation of Responses in Contingent Valuation Surveys", in P-O Johansson, B. Kriström and K-G Mäler, **Current Issues in Environmental Economics**, Manchester, Manchester University Press.

Huang, J-C, and V. K. Smith, 1998. "Monte Carlo Benchmarks for Discrete Response Valuation Methods", *Land Economics*, 74 (2 May): 186-202.

Kriström, B., 1990. "A Non-Parametric Approach to the Estimation of Welfare Measures in Discrete Response Valuation Studies", *Land Economics*, 66(2 May): 135-139.

Noether, G. E. 1967. **Elements of Nonparametric Statistics**. New York: Wiley.

Park, T., J. B. Loomis, and M. Creel, 1991. "Confidence Intervals for Evaluating Benefits Estimates from Dichotomous Choice Contingent Valuation Studies", *Land Economics*, 67(1 Feb):64-73.

Poe, G. L., E. K. Severance-Lossin, and M. P. Welsh, 1994. "Measuring the Difference (X-Y) of Simulated Distributions: A Convolutions Approach", *American Journal of Agricultural Economics*, 76(3 Nov): 904-915.

Ready, R. C., and D. Hu, 1995. "Statistical Approaches to the Fat Tail Problem for Dichotomous Choice Contingent Valuation", *Land Economics*, 71(4 Nov): 491-499

Ready, R. C., S. Navrud, and W. R. Dubourg, 2001. "How do Respondents with Uncertain Willingness to Pay Answer Contingent Valuation Questions?", *Land Economics* 77(3 Aug): forthcoming.

Welsh, M. P., and G. L. Poe, "Elicitation Effects in Contingent Valuation: Comparisons to a Multiple Bounded Discrete Choice Approach", *Journal of Environmental Economics and Management*, 36:170-185.

Zar, J. H., 1996. **Biostatistical Analysis**, Third Edition, London, Prentice Hall International, Inc.

Table 1: Results from the Dichotomous Choice Survey and Conversion of Open-Ended Responses to Dichotomous Choices

Price	Dichotomous Choice		Converted Open-Ended	
	Total Number of Observations	Percentage that Accepted the Posted Price	Average Number of Observations	Average Percentage that Would have Accepted the Posted Price
$1	94	93.62%	77	88.00%
$4	174	67.82%	143	55.92%
$6	31	32.26%	25	16.12%
$8	87	24.14%	71	8.43%
$12	35	11.43%	29	4.90%

Source: Taken from Table 3 in Balistreri et al., 2001.

Table 2: Comparison of Mean and Median WTP Estimates

Parametric Specifications, Assuming Non-Negativity	Mean	Std. Error	95% C. I.	Pr(= OE)[c]	PR(=EA)[d]
Log-logistic	7.57	1.1837	[6.34, 10.56]	<0.001	0.000[e]
Truncated Log-logistic	6.69	0.4008	[5.99, 7.55]	<0.001	0.000[e]
Pinched Log-logistic	6.36	0.4417	[5.70, 7.42]	<0.001	0.000[e]
Log-normal	7.25	0.8243	[6.15, 9.29]	<0.001	0.000[e]
Truncated Log-normal	6.86	0.4595	[6.03, 7.82]	<0.001	0.000[e]
Pinched Log-normal	6.27	0.4668	[5.62, 7.43]	<0.001	0.000[e]
Weibull	6.04	0.3558	[5.46, 6.85]	<0.001	0.000[e]
Truncated Weibull	6.04	0.3574	[5.46, 6.45]	<0.001	0.000[e]
Pinched Weibull	6.60	0.4713	[5.88, 7.71]	<0.001	0.000[e]
Non-Parametric Specifications					
Kriström	5.87	0.3150	[5.23, 6.48]	0.003	0.000[e]
Turnbull	4.56	0.3079	[3.96, 5.16]	0.961	0.009
Parametric Specifications, Allowing for Negativity.					
Linear Logistic					
Mean	5.63	0.2640	[5.14, 6.18]	0.006	0.000[e]
Non-Negative Mean[a]	5.77	0.2672	[5.30, 6.35]	0.002	0.000[e]
Linear Normal					
Mean	5.72	0.2644	[5.22, 6.26]	0.003	0.000[e]
Non-Negative Mean[b]	5.80	0.2679	[5.33, 6.38]	0.001	0.000[e]

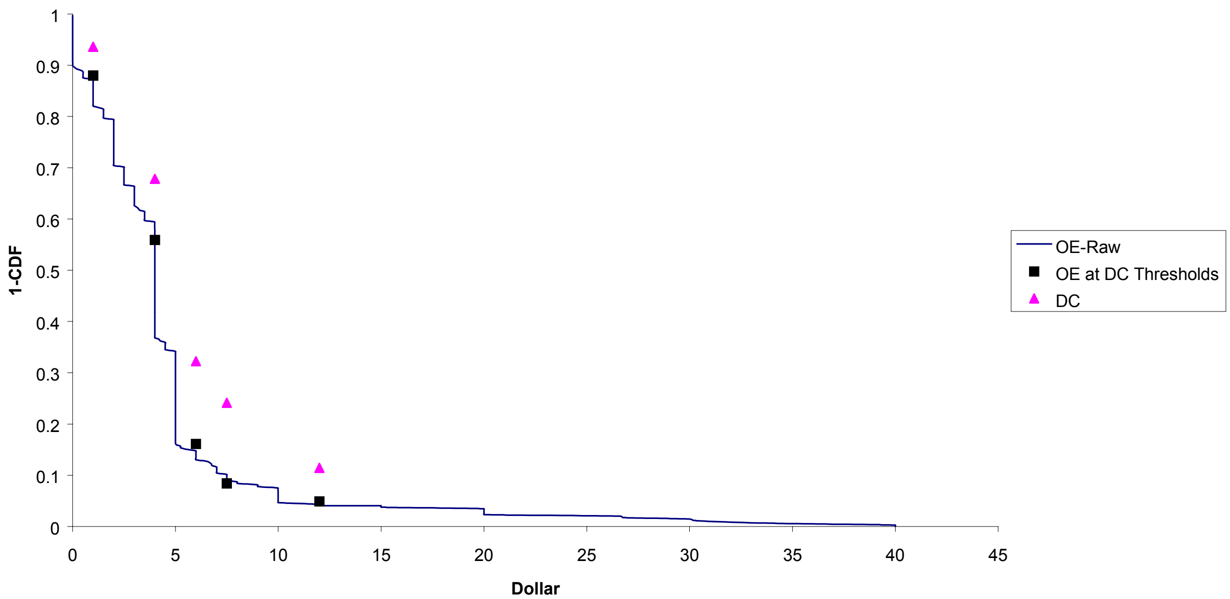
[a] Non-negative mean calculated using formula in HS footnote 13.
[b] Non-negative mean calculated using numerical integration.
[c] OE (open ended) values are mean=4.58, standard error=0.2894, 95% CI=[4.01, 5.15].
[d] EA (English Auction) values are mean=3.66, standard error=0.1595, 95% CI=[3.34, 3.97].
[e] The two vectors being compared do not overlap.

Table 3: Mean Willingness to Pay Estimates for Various Bid Vectors

Data Description [Bid Vector]	Turnbull	Kriström	Pinched Log-Logistic	Pinched Weibull	Pinched Log-Normal
Full Bid Vector [$1,$4,$6,$8,$12]	4.56	5.87	6.36	6.60	6.27
Jackknife Bid Vector [$1,$4,$6,$8]	4.10	5.67	6.26	6.25	6.06
Jackknife Bid Vector [$1,$4,$6,$12]	4.30	5.89	6.58	6.96	6.62
Jackknife Bid Vector [$1,$4,$8,$12]	4.39	6.15	6.52	6.77	6.41
Jackknife Bid Vector [$1,$6,$8,$12]	3.49	5.60	6.04	6.32	5.92
Jackknife Bid Vector [$4,$6,$8,$12]	4.30	5.84	6.24	6.57	6.19
Mean Squared Error[a]	0.30	0.04	0.04	0.07	0.06

[a] Using value for "Full Bid Vector" as the reference level.

Figure 1: Survival Functions (F($)): Raw Open-Ended, Open-Ended at DC Thresholds, and DC



Footnotes

---

[i] HS cite an earlier working paper by Balistreri et al. in their study. The difference between the two versions is largely editorial.

[ii] Throughout, two-tailed tests and 5 percent levels of significance are used.

[iii] The linear logistic function and the derivation of the mean WTP values from this function are provided in Footnote 13 in HS.

[iv] In making this conversion Balistreri et al. sought to maintain independence in the converted OE responses across prices. To accomplish this, each OE value was allocated randomly to one of the five prices in a way that produced sample sizes proportional to the DC samples for each price. It should be readily apparent that the results from such an exercise are dependent on the random allocation. To get a proportion at each price that was not dependent on a particular allocation, 100 random allocations were used; the average proportions from these 100 allocations are reported in the last column of Table 1.

[v] It is interesting to point out that no such irrationalities occurred in the actual money decisions made in the English Auction treatment. This may be attributed to either the fact that the realities of actual money insured rationality, or that the group auction mechanism used provided information to otherwise irrational respondents, or both.

[vi] Interestingly, while HS show that the specification errors lead to differential mean squared errors in the Monte Carlo simulations, they do not indicate the direction that any bias would take.

[vii] In the laboratory experiment, the DC respondents all received the same (hypothetical) endowment from which to purchase insurance against an expected loss; there is also no differentiation between nonuse and use values; and, the participants are undergraduate students and as such constitute a more or less homogenous population with similar relative prices and income levels. Hence, the Monte Carlo simulation results are irrelevant to our situation.

[viii] Conceptually, we realize that the Turnbull lower bound estimator for the open-ended responses is simply that associated with the "continuous" survival function provided in Figure 1. We use this term in the text simply to demonstrate our point.

[ix] In introducing the Balistreri et al. paper, HS assert that this "study is notable in considering the importance of bid design for the performance of the [dichotomous choice] approach" (p. 190).

OTHER A.E.M. WORKING PAPERS

WP No	Title	Fee (if applicable)	Author(s)
2001-17	Measuring the Effects of Eliminating State Trading Enterprises on the World Wheat Sector		Maeda, Koushi, N. Suzuki, and H. M. Kaiser
2001-16	Disentangling the Consequences of Direct Payment Schemes in Agriculture on Fixed Costs, Exit Decision and Output		Chau, N. H. and H. deGorter
2001-15	Market Access Rivalry and Eco-Labeling Standards: Are Eco-Labels Non-tariff Barriers in Disguise?		Basu, A. K. and N. H. Chau
2001-14	An Economic Evaluation of the New Agricultural Trade Negotiations: A Nonlinear Imperfectly Competitive Spatial Equilibrium Approach		Kaiser, H. M., N. Suzuki and K. Maeda
2001-13	Designing Nonpoint Source Pollution Policies with Limited Information about Both Risk Attitudes and Production Technology		Peterson, J. M. and R. N. Boisvert
2001-12	Supporting Public Goods with Voluntary Programs: Non-Residential Demand for Green Power		Fowlie, M., R. Wiser and D. Chapman
2001-11	Incentives, Inequality and the Allocation of Aid When Conditionality Doesn't Work: An Optimal Nonlinear Taxation Approach		Kanbur, R. and M. Tuomala
2001-10	Rural Poverty and the Landed Elite: South Asian Experience Revisited		Hirashima, S.
2001-09	Measuring the Shadow Economy		Kyle, S. and A. Warner
2001-08	On Obnoxious Markets		Kanbur, R.
2001-07	The Adoption of International Labor Standards		Chau, N. and R. Kanbur
2001-06	Income Enhancing and Risk Management Perspective of Marketing Practices		Peterson, H. and W. Tomek
2001-05	Qual-Quant: Qualitative and Quantitative Poverty Appraisal: Complementarities, Tensions and the Way Forward		Kanbur, R. (Editor)

Paper copies are being replaced by electronic Portable Document Files (PDFs). To request PDFs of AEM publications, write to (be sure to include your e-mail address): Publications, Department of Applied Economics and Management, Warren Hall, Cornell University, Ithaca, NY 14853-7801. If a fee is indicated, please include a check or money order made payable to Cornell University for the amount of your purchase. Visit our Web site (*http://aem.cornell.edu/research/workpaper3.html)* for a more complete list of recent bulletins.