Pinuccia Calia
University of Cagliari
calia@unica.it

Elisabetta Strazzera
University of Cagliari
strazzera@unica.it

BIAS AND EFFICIENCY OF SINGLE VS. DOUBLE BOUND MODELS FOR CONTINGENT VALUATION STUDIES: A MONTE CARLO ANALYSIS*

Abstract: The *Dichotomous Choice Contingent Valuation Method* (DC-CVM), both in the single and the double bound formulation, has been in the last years the most popular technique among practitioners of contingent valuation, due to its simplicity of use in data collection. The *single bound* procedure is easier to implement than the *double bound*, especially in data collection and estimation. On the other hand, it is well known that the double bound is more efficient than the single bound estimator. It remains to analyze the bias of the ML estimates produced by either model, and the gain in efficiency associated to the double bound model, in different experimental settings. We find that there are no relevant differences in point estimates given by the two models, even for small sample size, so that neither estimator can be said to be less biased than the other. The greater efficiency of the double bound is confirmed, although it can be seen that the differences tend to reduce by increasing the sample size, and are often negligible for medium size samples. Provided that a reliable *pre-test* is conducted, and the sample size is large, our results warrant the use of the single rather than the double bound model.

* We would like to thank Rossella Diana for her helpful assistance in the early stage of this research.

1. Introduction

The *Dichotomous Choice Contingent Valuation Method* (DC-CVM) has been in the last years the most popular technique among practitioners of contingent valuation, due to its simplicity of use in data collection. When this elicitation method is used, the respondent is only required to answer YES or NO when asked if she/he is willing to pay a given amount (*bid*) for the public good. The *single bound* model comprises only one such question, while in the *double bound* model the first question is followed by another specifying a lower amount, if the answer to the first question was negative, and higher otherwise. This procedure is certainly easier for respondents than other methods requiring long adjustment processes, like the bidding game; or a precise assessment of the individual's own reservation price based on introspective analysis, as it happens in the open ended elicitation method. The price to pay for this is the limited information arising from DC-CVM data: the only information available to the researcher after the interview is an interval of values containing the true willingness to pay (*wtp*) of the individual.

In the single bound model the interval is bounded by the bid and the limit of the *wtp* distribution (the upper limit if the answer was positive, the lower limit otherwise). In the double bound model the interval is enclosed within two bids, if one answer to the two questions was positive and the other negative (double bound); otherwise, the interval is bounded by the second bid and the limit of the *wtp* distribution. In order to gather more information about the support of the true *wtp* distribution, the initial bids are varied among individuals.

Hanemann, Loomis and Kanninen (1991) proved that the double bound DC-CVM is asymptotically more efficient than the single bound model; empirical results by these authors, and by Leòn (1995), confirm this property also for finite samples. These studies show that point estimates for the mean and median *wtp* produced by the two models are substantially different. Some authors interpret this finding arguing that the double bound model produces not only more efficient but also less biased estimates than the single bound. Hanemann *et al*. (1991), for example, suggest that the double bound model allows for correction of a poor choice of the initial vector of bids. The same contention is also purported by Kanninen (1995): with real data and assuming that the *wtp* distribution of the population is a Logistic, she calculates the bias of the double and the single model estimates, finding out that the latter is larger. This can be hardly thought to be a definitive answer, though, given the small sample (100 observations) considered in her study, and, more fundamentally, that her assumption about the true *wtp* model might have been incorrect.[1] To shed some light on the matter, we carry on a Monte Carlo study, designed to compare the performance of either estimator under different experimental situations. After a quick overview of the two models (section 2), we present in the following sections the experimental setting (section 3) and the results (section 4); section 5 concludes the paper.

2. The wtp models

We adopt the censored econometric model proposed by Cameron and James (1987) and Cameron (1988), which, unlike the utility differential model of Hanemann (1984), produces separate estimates for the standard deviation of the *wtp* and the parameters of the model. This allows us to easily compute the confidence intervals for the central tendency measures of *wtp*: as described later, estimates of the standard errors of the coefficients are directly plugged in an analytical formula. It is worth to mention that only recently confidence intervals (either derived through analytical calculus or through bootstrapping methods[2]) for the *wtp* estimate are being included in contingent valuation studies[3].

Assuming a linear functional form for the *wtp*, the econometric model is the following:

(2.1) $$Y_i = x_i' b + e_i$$

where Y_i is the true individual willingness to pay, which is assumed to depend on individual socioeconomic characteristics contained in the vector x_i. The error term e_i is distributed with c.d.f. $F(e_i)$ with zero mean and variance equal to v^2. In this model Y_i is considered a latent continuous censored variable: the observed variable is the answer YES or NO to the question regarding whether or not the individual would be willing to pay a given amount t_i. Letting P_1 the probability that the reservation price Y_i for a given individual is greater than t_i, and P_0 the complementary probability, the single bound model is specified as follows:

---

[1] Another interpretation is that the answers to the first question and the follow up come from two different distributions, even though correlated: cfr. Herriges and Shrogren (1996), and Cameron and Quiggin (1994). Other sources of disturbance on the data arising from the follow up question are analyzed by Alberini, Kanninen and Carson (1997). For a comparison between the univariate and the bivariate double bound model cfr. Alberini (1995).

[2] For a comparison between different methods cfr. Cooper (1994).

[3] Cfr. Cameron (1991) and Park, Loomis and Creel (1991).

(2.2)
$$\begin{aligned} P_1 &= \Pr(Y_i > t_i) \\ &= \Pr(x_i' b + e_i > t_i) \\ &= \Pr(e_i > t_i - x_i' b) \end{aligned}$$

and, after standardization,

(2.3) $P_1 = \Pr(z_i > (t_i - x_i' b)/v)$.

For a given sample of n independent observation, the log-likelihood function is:

(2.4)
$$LogL = \sum_{i=1}^{n} \{ I_i \log [1 - F((t_i - x_i' b)/v)] + (1 - I_i) \log [F((t_i - x_i' b)/v)] \},$$

where I_i is a dummy variable assuming value one if the answer is positive, zero otherwise.

Since $1/v$ is the coefficient of the bid t_i and bids are varied among individuals, b and v can be estimated separately, so we have a direct estimate of the standard deviation of *wtp*.

When the double bound model is chosen instead, we observe two dichotomous variables, i.e. the answers to the first question and its follow up. This method produces four possible outcomes, with probabilities as follows:

(2.5)
$$\begin{aligned} \Pr(yes, yes) &= \Pr(Y_i \geq t_i \geq t_i^u) = 1 - F(t_i^u) \\ \Pr(yes, no) &= \Pr(t_i \leq Y_i \leq t_i^u) = F(t_i^u) - F(t_i) \\ \Pr(no, yes) &= \Pr(t_i^l \leq Y_i \leq t_i) = F(t_i) - F(t_i^l) \\ \Pr(no, no) &= \Pr(Y_i \leq t_i^l \leq t_i) = F(t_i^l) \end{aligned}$$

with log-likelihood function:

$$LogL = \sum_{i=1}^{n} \Big\{ I_i \; I_i^u \; \log \left[F\left((t_i^u - x_i' b)/v \right) \right] +$$

(2.6)
$$+ I_i \left(1 - I_i^u\right) \log \left[F\left((t_i^u - x_i' b)/v \right) - F\left((t_i - x_i' b)/v \right) \right] +$$
$$+ I_i^l \left(1 - I_i\right) \log \left[F\left((t_i - x_i' b)/v \right) - F\left((t_i^l - x_i' b)/v \right) \right] +$$
$$+ \left(1 - I_i\right) \left(1 - I_i^l\right) \log \left[F\left((t_i^l - x_i' b)/v \right) \right] \Big\}$$

Here t_i stays for the bid offered in the first question; t_i^u is the follow up if the answer to the first question has been positive; t_i^l is the follow up when the answer to the first question has been negative. I_i, I_i^u, I_i^l are dichotomous variables with value one if the answer to the first bid or the corresponding follow-up has been positive, and zero otherwise.

Once the parameters of either model are estimated, through Maximum Likelihood procedure, estimation of the mean *wtp* is straightforward: it suffices to calculate

(2.7) $$E(Y) = \bar{x}'\hat{b}$$

where $\bar{x}$ is the vector of sample averages of the regressors and $\hat{b}$ is the vector of ML estimates of the parameters. Another measure of interest in contingent valuation studies, especially when the *wtp* distribution is asymmetric, is the estimate of the median *wtp*, whose analytical form depends on the *wtp* distribution.

It is useful to calculate also confidence intervals for the mean or median *wtp*. Only recently researchers have begun to include confidence intervals in their reported fitted *wtp* measures, either using refinements of the bootstrap method (Krinsky and Robb (1986); McLeod and Bergland (1989)) or using the analytical formula proposed by Cameron (1991). For the model in eq. (2.1), Cameron demonstrated that an interval for $E(Y)$ at significance level a can be calculated as follows:

(2.8) $$CI_{1-a}\left[E(Y)\right] = \bar{x}'\hat{b} \pm t_{a/2} \sqrt{\bar{x}' S_b \bar{x}}$$

where S_b is an estimate of the variance-covariance matrix of the parameter estimates. In a paper by Cooper (1994) it is shown that either method to calculate confidence intervals performs quite well, the relative ranking depending on sample size and specification of the *wtp* model. Given the simplicity of Cameron's method, we use her analytical formula to calculate confidence intervals for the fitted mean (median) values of *wtp* in our experiments.

3. The Monte Carlo Study

We consider two specifications for the *wtp* among the most commonly used in CVM studies. The first one is a linear equation for the latent variable:

(3.1)

The second specification is a logarithmic function:

(3.2)

The variables e and h are error terms with zero mean and variance s^2 and t^2 respectively. In designing the *Monte Carlo* analysis we assume, for specification (3.1), that *wtp* has two different distributions with mean and variance equal to s^2: the first is Normal, the second is a mixture of two Normal which resembles an asymmetric distribution.

For specification (3.2), *wtp* is assumed to have a lognormal distribution (so that the error term h has a Normal distribution) with mean , median
and variance:

.

This specification is particularly suited to account for asymmetries in the *wtp* distribution, often observed in real data.

For each specification we generate 200 samples with four different size: 100, 250, 400 and 1000 observations.

3.1. The linear specification

The *wtp* data is generated according to the model:

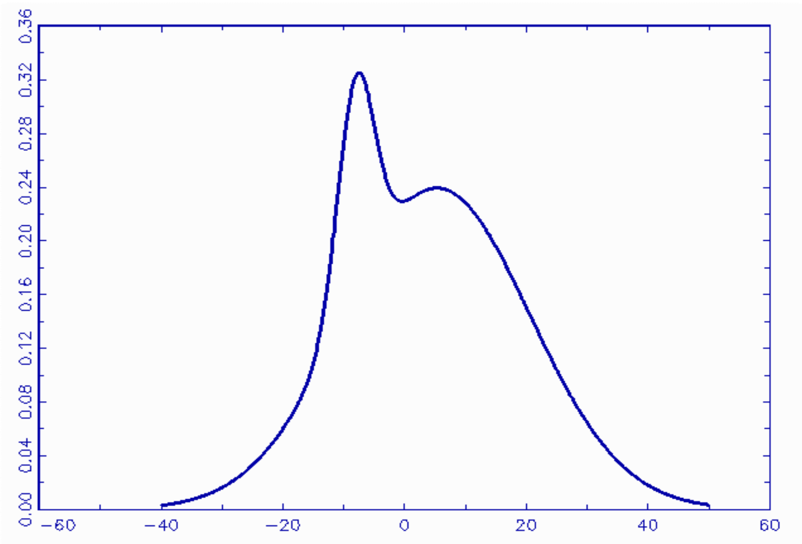
(3.1.1)

with $a = 20$, $b = 0.1$ and values of the regressor x drawn from a Uniform distribution in the range 40-750. In a first set of experiments, the error term e is a Normal variable with zero mean and standard deviation $s = 10$. In another experiment, we consider a situation where the error term has a mixture distribution:

(3.1.2)

where and . Varing the values of p, m_1, m_2, s_1^2 and s_2^2, $f(e\ )$ is allowed to assume different forms (either simmetric or asymmetric, unimodal or bimodal). We choose $p = 0.4$, $m_1 = -8$, $m_2 = 5.33$, $s_1 = 3$ and $s_2 = 15.22$, in order e to have mean equal to zero and standard deviation equal to 10; besides, the resulting distribution is somehow bimodal and asymmetric with a heavy right tail (Figure 1).

Fig. 1. Density function of the mixture distribution



We report results for two bid designs: in the first (bid design A), the chosen bids are the quartiles of the *wtp* empirical distribution in a small independent sample (50 observations), simulated through eq. (3.1.1) with no error term. In the second experiment instead (bid design B), three bid values (10, 20, 30) are selected such that only the left tail of the *wtp* distribution is covered (less than 15% for both *wtp* distributions).

In both experiments, the selected values are then randomly assigned to the individuals of the sample and compared with the corresponding Y_i in order to create the dichotomous variable for the first answer. The follow ups required for the double bound model are obtained from the first bid by increasing or decreasing it by 25% of its amount: whenever the first bid is lower than Y_i, the bid is reduced; otherwise, it is increased. The dichotomous dependent variable, I_i assumes value zero if the true *wtp* is lower or equal to the assigned bid; otherwise it assumes value one. For the double bound model, we have two dependent variables: the first is generated, as before, by comparing each Y_i to the assigned first bid; the second is obtained analogously, matching Y_i with the second bid.

The two models are estimated with ML procedure, using the log-likelihood functions (2.4) and (2.6), respectively for the single and the double bound model. The Normal c.d.f. *F* is plugged in the log-likelihood functions in place of the generic c.d.f. *F*: the model is therefore correctly specified in both the deterministic and the stochastic part when the *wtp* distribution is normal, while we are allowing for misspecification in the stochastic part of the model when the *wtp* distribution is asymmetric.

3.2. The loglinear specification

The *wtp* data for the loglinear model is simulated according to the following equation:

(3.2.1)

where l = 1.05, d = 0.35 and x is a Uniform regressor in the range 2500-125000[4].

The disturbance h is simulated from a Normal distribution with mean zero and standard deviation t = 1.48.

The bids are selected as percentiles (5th, 10th, 20th, 45th, 75th, 95th) of the *wtp* (obtained as exp(ln (Y))) empirical distribution in a small independent sample. Analogously to the experiment with the linear model, the follow-up bid is created by increasing, or decreasing, the first bid by 25% of its amount. The dichotomous dependent variables are then created by comparing ln(Y_i) to the logarithm of the first bid and, sequentially, to that of the appropriate follow up.

For this experiment we assume a correct specification of the model, so that for estimation of the single and the double bound models the normal c.d.f. *F* is substituted for *F* in the log-likelihood equation (2.4) and (2.6) respectively.

Given the asymmetric shape of the *wtp* distribution generated by eq. (3.2.1), the median rather than the mean value can be indicated as an appropriate measure of central tendency. In such a case the calculus of the confidence intervals follows two steps: in the first step we calculate the limits of the interval around E (ln(Y)); then, we transform these values by taking the anti-log. This is a correct confidence interval for the median (cfr. Greene (1991, pag.168) for OLS estimates and Cameron (1991) for ML estimates of the loglinear model parameters); the results,

---

[4] These values for the parameters and the regressor are taken from Jordan and Elnagheeb (1994).

though, are not entirely satisfactorily, as it will be seen in the next section.

4. Results

Results for all *Monte Carlo* experiments are reported in Tables 1 through 5.

The theoretical result about the higher efficiency of the double bound model is confirmed in all experiments: the standard deviations of the estimates from the double bound are always smaller than those obtained from the single bound. Point estimates from both models get more precise when the number of observations increases, though, and for medium size samples the differences in efficiency are often negligible.

The results about the bias of the estimates obtained from the models instead are not so clear cut. Although the double bound model generally gives less biased parameter estimates, this is not always true for the estimate of the standard deviation of *wtp*. Furthermore, the central tendency measures are in some cases estimated more accurately by the single bound model, even though the opposite holds more often. Yet, as we can see from the results in the following tables, there are no substantial differences in bias for the relevant measure of *wtp* between the two models.

More remarkable differences can be found in the estimates of confidence intervals: as it can be expected, the double bound model gives narrower intervals (about half the length of corresponding interval of the single model). As a consequence of this, and since the bias of the estimated mean or median *wtp* for the two models is quite similar, the double bound model produces also intervals with lower empirical confidence level in almost all experiments. It can be noticed that in general, for small sample size, the estimated confidence intervals are not much reliable: for acceptable interval lengths the empirical level is generally far away from the nominal confidence level, while empirical levels closer to the nominal are associated to wide intervals. This problem reduces for higher sample dimensions, where we find narrower intervals and empirical levels closer to the nominal confidence level of 90%.

Table 1(a). Linear model (bid design A): average and standard deviation (in parenthesis) of parameter point estimates across 200 replications

Estimates	Sample size			
	100	**250**	**400**	**1000**
α (20)				
Single	19.471	19.754	19.936	19.925
	(6.469)	(4.105)	(3.010)	(2.047)
Double	19.878	19.890	19.986	20.044
	(3.531)	(2.307)	(1.848)	(1.209)
β (0.1)				
Single	0.101	0.100	0.100	0.100
	(0.0104)	(0.009)	(0.007)	(0.005)
Double	0.100	0.100	0.100	0.100
	(0.007)	(0.005)	(0.004)	(0.003)
σ (10)				
Single	9.699	9.843	10.059	9.976
	(2.338)	(1.487)	(1.112)	(0.748)
Double	9.971	9.869	10.019	9.965
	(1.397)	(0.778)	(0.603)	(0.377)

Table 1(b). Linear model (bid design A): summary statistics on estimated mean *wtp* across 200 replications

Estimates	Sample size			
	100	**250**	**400**	**1000**
E(Y) (59.5)				
Single	63.184 [a]	61.014	60.687	59.820
	(2.255) [b]	(1.181)	(1.034)	(0.641)
Double	63.228	61.002	60.649	59.860
	(1.476)	(0.842)	(0.663)	(0.473)
Bias (E(Y))				
Single	3.684	1.513	1.187	0.320
Double	3.727	1.501	1.149	0.360
Conf. Level				
Single	44.7 [c]	64.5	63.5	85.5
Double	18.1	47.0	48.5	78.0
Average width				
Single	6.976 [d]	4.211	3.335	2.115
Double	4.629	2.868	2.280	1.442

[a] Average of estimated mean *wtp*; [b]Standard deviation of estimated mean *wtp*; [c]Empirical confidence levels: percentage of inclusion of true mean *wtp* into the confidence intervals; [d]Mean difference between upper and lower limits

A comparison of tables 1(a) and 1(b) with tables 2(a) and 2(b) shows that a wrong bid design (design B) affects to some extent the performance of both models: estimates are more biased and less precise, in particular for small sample size. Especially severe is the increase in the standard deviation of mean *wtp* estimates, which is reflected also in the marked increase, for small sample size, of the width of the confidence intervals.

Table 2(a). Linear model (bid design B): average and standard deviation (in parenthesis) of parameter point estimates across 200 replications

Estimates	Sample size			
	100*	**250**	**400**	**1000**
α (20)				
Single	19.047	19.651	19.767	19.825
	(6.396)	(4.381)	(3.023)	(1.923)
Double	19.219	19.855	19.849	19.910
	(3.990)	(2.423)	(2.057)	(1.224)
β (0.1)				
Single	0.104	0.102	0.101	0.101
	(0.029)	(0.018)	(0.012)	(0.007)
Double	0.103	0.100	0.100	0.100
	(0.014)	(0.008)	(0.007)	(0.004)
σ (10)				
Single	8.355	9.643	9.767	9.931
	(2.893)	(1.963)	(1.556)	(0.848)
Double	9.450	9.774	9.828	9.967
	(1.725)	(1.177)	(0.922)	(0.514)

*Two replications giving abnormal values have been dropped off from the results of the single bound model

Table 2(b). Linear model (bid design B): summary statistics on estimated mean *wtp* across 200 replications

Estimates	Sample size			
	100*	**250**	**400**	**1000**
E(Y) (59.5)				
Single	63.914 [a]	61.609	60.840	59.982
	(7.806) [b]	(3.776)	(2.461)	(1.516)
Double	63.648	61.114	60.731	59.898
	(3.545)	(1.850)	(1.440)	(0.899)
Bias (E(Y))				
Single	4.414	2.109	1.339	0.482
Double	4.148	1.614	1.231	0.398
Conf. Level				
Single	91.7 [c]	87.0	94.0	91.5
Double	76.5	81.5	84.0	88.0
Average width				
Single	22.453 [d]	11.795	8.830	5.363
Double	11.162	6.071	4.682	2.926

*Two replications giving abnormal values have been dropped off from the results of the single bound model; [a] Average of estimated mean *wtp*; [b]Standard deviation of estimated mean *wtp*; [c]Empirical confidence levels: percentage of inclusion of true mean *wtp* in the confidence intervals; [d]Mean difference between upper and lower limits

Tables 3(a) and 3(b) show the results from the experiment where we consider a possible misspecification of the econometric model: we assume that the *wtp* is normally distributed while instead it is not.

It can be noticed that the two models are quite robust to misspecification, giving, in general, good point estimates of the parameters and mean *wtp*. The exception is the estimate of the standard deviation of *wtp*, which is always overestimated by both models for all sample sizes. Anyway, comparison with the results reported in table 1(a) and 1(b) shows that misspecification affects in particular the precision of estimates, resulting in higher standard deviations.

Table 3(a). Linear model (bid design A, asymmetric mixture distribution): average and standard deviation (in parenthesis) of parameter point estimates across 200 replications

Estimates	Sample size			
	100	**250**	**400**	**1000**
α (20)				
Single	19.339	19.326	19.099	19.154
	(8.193)	(4.910)	(3.884)	(2.424)
Double	19.097	19.160	19.220	19.225
	(4.620)	(2.897)	(2.292)	(1.419)
β (0.1)				
Single	0.102	0.101	0.102	0.102
	(0.019)	(0.011)	(0.009)	(0.006)
Double	0.102	0.101	0.102	0.101
	(0.010)	(0.006)	(0.005)	(0.003)
σ (10)				
Single	12.726	13.381	13.484	13.618
	(3.217)	(1.963)	(1.539)	(0.989)
Double	13.058	13.408	13.447	13.504
	(1.623)	(1.024)	(0.810)	(0.518)

Table 3(b). Linear model (bid design A, asymmetric mixture distribution): summary statistics on estimated mean *wtp* across 200 replications

Estimates	Sample size			
	100	**250**	**400**	**1000**
E(Y) (59.5)				
Single	63.491[a]	60.974	60.609	59.924
	(2.641) [b]	(1.615)	(1.210)	(0.760)
Double	63.283	60.821	60.560	59.722
	(1.901)	(1.118)	(0.884)	(0.548)
Bias (E(Y))				
Single	3.991	1.474	1.109	0.424
Double	3.783	1.321	1.060	0.222
Conf. Level				
Single	51.5 [c]	73.5	77.5	85.5
Double	34.0	66.5	66.0	88.0
Average width				
Single	8.187 [d]	5.066	4.005	2.559
Double	5.820	3.668	2.898	1.843

[a] Average of estimated mean *wtp;* [b]Standard deviation of estimated mean *wtp*; [c]Empirical confidence levels: percentage of inclusion of true mean *wtp* in the confidence intervals;
[d]Mean difference between upper and lower limits

In tables 4(a) and 4(b) are reported the results of the experiment with the asymmetric distribution and bid design B. When misspecification and bad bid design combine, the optimization algorithm fails to converge several times, particularly for small sample size. Also, for the same sample size, we found that abnormal values for point estimates of the parameters are produced in many replications by the single bound model, while the double bound is more robust. In calculating the summary statistics, the replications with such abnormal values are dropped from the sample.

Table 4(a). Linear model (bid design B, asymmetric mixture distribution): average and standard deviation (in parenthesis) of parameter point estimates across 200 replications

Estimates *	Sample size			
	100 [i]	**250** [ii]	**400** [iii]	**1000** [iv]
α (20)				
Single	17.316	17.283	17.516	17.858
	(14.814)	(5.004)	(3.540)	(2.038)
Double	16.809	17.918	17.811	18.036
	(6.049)	(3.371)	(2.677)	(1.582)
β (0.1)				
Single	0.119	0.107	0.013	0.098
	(0.139)	(0.042)	(0.029)	(0.016)
Double	0.116	0.102	0.103	0.101
	(0.041)	(0.020)	(0.016)	(0.009)
σ (10)				
Single	10.045	9.734	9.729	9.583
	(12.156)	(3.512)	(2.478)	(1.411)
Double	10.313	9.945	10.362	10.159
	(3.015)	(1.691)	(1.042)	(0.821)

[i] 21 replications are deleted because of failure to convergence and 14 for the single bound and 3 for the double bound model due to abnormal parameter values; [ii] 3 replications are deleted because of failure to convergence and 5 for the single bound model due to abnormal parameter values; [iii] 1 replication is deleted because of failure to convergence and 2 for the single bound and 1 for the double bound model due to abnormal parameter values; [iv] 1 replication for the single bound model is deleted due to abnormal parameter values

Table 4(b). Linear model (bid design B, asymmetric mixture distribution): summary statistics on estimated mean *wtp* across 200 replications

Estimates	Sample size			
	100 [i]	**250** [ii]	**400** [iii]	**1000** [iv]
E(Y) (59.5)				
Single	69.038 [a]	61.307	59.490	57.063
	(37.742) [b]	(16.092)	(9.238)	(5.400)
Double	67.172	59.691	59.938	58.168
	(14.870)	(6.172)	(4.715)	(2.545)
Bias (E(Y))				
Single	9.537	1.807	-0.010	-2.437
Double	7.672	0.191	0.438	-1.332
Conf. Level				
Single	96.3 [c]	85.9	86.8	74.4
Double	93.2	89.3	89.8	84.5
Average width				
Single	163.55 [d]	44.83	30.44	16.41
Double	44.36	19.34	15.46	8.78

[i]21 replications are deleted because of failure to convergence and 14 for the single bound and 3 for the double bound model due to abnormal parameter values; [ii]3 replications are deleted because of failure to convergence and 5 for the single bound model due to abnormal parameter values; [iii]1 replication is deleted because of failure to convergence and 2 for the single bound and 1 for the double bound model due to abnormal parameter values; [iv]1 replication for the single bound model is deleted due to abnormal parameter values; [a] Average of estimated mean *wtp;* [b]Standard deviation of estimated mean *wtp*; [c]Empirical confidence levels: percentage of inclusion of true mean *wtp* in the confidence intervals; [d]Mean difference between upper and lower limits

It is quite clear that for this experimental design, the double bound performs better. Especially for small sample size, the double bound secures a relevant gain in efficiency, although, as usual, the differences tend to decrease when working with more observations.

This can be noticed also by looking at the average width of the confidence intervals: the proportion of the single bound interval width with respect to the corresponding double bound interval is about 3.5 for the sample size of one hundred and falls to 2 for greater sample sizes. It is also interesting that for this experiment design the confidence levels associated to the double bound intervals are always better than the single bound.

Differences in bias instead are not so significant. Taking into account the misspecification and the very poor bid design, we can say that both estimators perform reasonably well in giving point estimates for the parameters and of the mean wtp, at least for sample sizes 250 and over. It seems worth to point out that the irregular pattern of the two estimators performance across sample sizes can also be attributed to sampling variability introduced by dropping a different number of replications in each experiment.

Finally, tables 5(a) and 5(b) report the results of the experiment with the loglinear specification.

In spite of the correct specification and good bid design (such that most of the *wtp* distribution is covered) we observe large bias and standard deviation values, in particular if compared with the analogous experimental design for the linear model. Some problems arise especially in estimating the parameter λ, for which negative values are produced in many replications, and whose bias is more severe than that of the other parameters. As usual, the more the observations the better the estimates.

Our application of Cameron's analytical formula to the loglinear model is not very satisfying: the values of the average width cast some doubts about the appropriateness of the procedure adopted, described in section 3. But this result also is in line with Cooper's (1994) finding that when the distribution is asymmetric Cameron's technique is not much reliable, and bootstrap methods for calculating confidence intervals should be preferred.

Table 5(a). Loglinear model: average and standard deviation (in parenthesis) of parameter point estimates across 200 replications

Estimates	Sample size			
	100	**250**	**400**	**1000**
λ (1.05)				
Single	0.718	0.813	0.850	0.952
	(3.129)	(2.039)	(1.427)	(0.972)
Double	1.020	0.991	1.009	0.991
	(2.417)	(1.573)	(1.189)	(0.808)
δ (0.35)				
Single	0.379	0.370	0.368	0.358
	(0.285)	(0.186)	(0.131)	(0.089)
Double	0.352	0.354	0.354	0.355
	(0.220)	(0.144)	(0.109)	(0.074)
τ (1.48)				
Single	1.439	1.471	1.482	1.473
	(0.288)	(0.168)	(0.126)	(0.082)
Double	1.467	1.467	1.485	1.475
	(0.189)	(0.110)	(0.081)	(0.051)

Table 5(b). Loglinear model: summary statistics on estimated median *wtp* across 200 replications

Estimates	Sample size			
	100	**250**	**400**	**1000**
M(Y) (125.87)				
Single	132.525 [a]	127.850	128.401	126.155
	(31.230) [b]	(16.490)	(14.632)	(9.307)
Double	132.231	128.283	128.944	126.194
	(24.730)	(13.170)	(11.026)	(7.157)
Bias (M(Y))				
Single	6.655	1.980	2.531	0.285
Double	6.361	2.413	3.074	0.234
Conf. Level				
Single	89.0 [c]	94.9	89.4	87.5
Double	84.0	92.4	88.4	88.5
Average width				
Single	100.770 [d]	60.212	47.809	29.310
Double	77.553	46.668	37.379	22.898

[a] Average of estimated median *wtp*; [b]Standard deviation of estimated median *wtp*; [c]Empirical confidence levels: percentage of inclusion of true median *wtp* in the confidence intervals;
[d]Mean difference between upper and lower limits

5. Conclusion

The single bound method presents some attractive features with respect to the double bound. It requires less information, it is easier to implement at data collection and estimation stages, and can avoid systematic bias in responses that are due to the introduction of the follow-up (for example, the so called "anchoring effect"). On the other hand, it is well known that the double bound is more efficient than the single bound estimator. It is therefore interesting to compare their behavior in terms of bias of the ML estimates produced by either model, and to analyze the gain in efficiency associated to the double bound model, in different experimental settings.

Our results confirm the theoretical findings about the efficiency of the double bound model. It produces more precise point estimates of parameters and central tendency measures of *wtp*, as well as narrower confidence intervals around mean or median *wtp*. The differences though tend to reduce by increasing the sample size, and are often negligible for medium size samples. On the contrary, no relevant differences can be found in point estimates given by the two models, even for small sample size, so that neither estimator can be said to be less biased than the other.

Granted that no other sources of systematic bias arise, and the sample size is large enough, huge differences in point estimates between the two models observed in some applications should probably be ascribed to misspecification of the model, or poor bid design, or, more probably, both. Generally, Contingent Valuation surveys are preceded by a *pre-test* survey on a small population sample, that allows to gather information about the *wtp* distribution. If the pre-test is conducted correctly, it gives a good *a priori* for the bid design of the survey; in such a case, use of the single bound model should be warranted. If instead the sample size is small, or the *pre-test* survey is not much reliable, it is advisable to use the double bound model: in these circumstances the gain in efficiency is so large that indeed may overwhelm other possible costs associated to the use of the double bound.

References

Alberini, A. (1995), “Efficiency vs Bias of Willingness to Pay Estimates: Bivariate and Interval-Data Models” *Journal of Environmental Economics and Management*, 29, pp.169-80.

Alberini, A., B.J. Kanninen and **R.T. Carson** (1997), “Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data”, *Land Economics*, forthcoming.

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

Cameron, T. A. (1991), “Interval Estimates of Non-Market Resource Values from Referendum Contingent Valuation Surveys”, *Land Economics*, 67, pp. 413-21.

Cameron, T. A. and **M.D. James** (1987), “Efficient Estimation Methods for Closed-Ended Contingent Valuation Surveys”, *The Review of Economics and Statistics*, 69, pp. 269-76.

Cameron, T. A. and **J. Quiggin** (1994), “Estimation Using Contingent Valuation Data from a ‘Dichotomous Choice with Follow-up’ Questionnaire”, *Journal of Environmental Economics and Management*, 27, pp. 218-34.

Cooper, J. (1994), “A Comparison of Approaches to Calculating Confidence Intervals for Benefit Measures from Dichotomous Choice Contingent Valuation Surveys”, *Land Economics*, 70, pp. 111-122.

Green, W. H. (1991), *Econometric Analysis*, MacMillan Publishing Company.

Hanemann, W.M. (1984), “Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses”, *American Journal of Agricultural Economics*, 66, pp. 332-41.

Hanemann, W. M., J. Loomis, and **B.J. Kanninen** (1991), “Statistical Efficiency of Double-bounded Dichotomous Choice Contingent Valuation”, *American Journal of Agricultural Economics*, 73, pp.1255-63.

Herriges, J.A. and **J.F. Shogren** (1994), “Starting Point Bias in Dichotomous Choice Valuation with Follow-up Questioning”, *Journal of Environmental Economics and Management*, 30, pp.112-31.

Kanninen, B.J. (1995), “Bias in Discrete response Contingent Valuation”, *Journal of Environmental Economics and Management*, 28, pp.114-25.

Krinsky, I. and **A. Robb** (1986), “Approximating the Statistical Properties of Elasticities”, *Review of Economics and Statistics*, 68, pp.715-19.

Jordan, G.L. and **A.H. Elnagheeb** (1994), "Consequences of Using Different Question Formats in Contingent Valuation: A Monte Carlo Study", *Land Economics*, 70, pp.97-110.

Leòn, C. J. (1995), “El metodo dicotomico de valoraciòn contingente: una aplicaciòn a los espacios naturales en Gran Canaria", *Investigaciones Economicas*, 19.

McLeod, D. and **O. Bergland** (1989), “The Use of Bootstrapping in Contingent Valuation Studies”, W.P., Dpt. of Agricultural Economics, Oregon State University, Corvallis

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

Finito di stampare nell'aprile 1998 presso la litografia SOLTER
Via Tolmino 33, Cagliari - Tel 282249

Impianti forniti da CRENoS