The Australian Journal of Agricultural and Resource Economics, 44:2, pp. 233–252

Dichotomous choice contingent valuation probability distributions†

Geoffrey N. Kerr*

Parametric distributions applied to dichotomous choice contingent valuation data invoke assumptions about the distribution of willingness to pay that may contravene economic theory. This article develops and applies distributions that allow the shape of bid distributions to vary. Alternative distributions provide little, if any, improvement in statistical fit from commonly used distributions. While median willingness to pay is largely invariant to distribution, estimates of mean consumer surplus diverge widely. Sensitivity analysis to determine benefit measure response to distributional assumptions is essential to prevent erroneous policy advice from applied dichotomous choice research.

1. Introduction

Dichotomous choice contingent valuation data are usually analysed by fitting parametric distributions[1] to the data to depict a representative individual's demand for the non-market good. In this way, the analyst has the ability to define the general shape of the function, and uses statistical analysis to identify the best-fitting function of the specified shape. Constraining the distribution of bids to a particular shape may bias welfare measure estimates when the fitted distribution does not closely resemble the underlying distribution of willingness to pay (WTP). An approach to avoiding this problem is the use of non-parametric, or semi-parametric, statistical methods for fitting response distributions (Carson *et al.* 1994a; Carson, Wilks and Imber 1994b; Creel and Loomis 1997; Haab and McConnell 1997; Kriström 1990; McFadden 1994). Results for these non-parametric and semi-parametric approaches are not substantially different from those obtained from parametric models.

† The author gratefully acknowledges the enthusiastic support and helpful comments of Ian Langford, Ian Bateman and Dominic Moran. Three anonymous referees provided extremely helpful suggestions that have greatly enhanced the quality of the final product.

* Geoffrey N. Kerr, Environmental Management and Design Division, Lincoln University, New Zealand.

[1] Referred to by a variety of names, including (*inter alia*): bid function, link function, probability distribution, cumulative density function, and survival function.

The most commonly utilised parametric distribution is the logistic function, using either raw or logged bid amounts.[2] The normal distribution and its logged variant have also been used, but because the results are virtually identical to the logistic models (Bowker and Stoll 1988), the logistic models are generally preferred for their mathematical tractability. Other distributions that have been utilised include Hanemann's (1984) variant of the logistic model, which includes income effects, McFadden's (1994) flexible model, which also includes income effects, and the Weibull distribution (Carson *et al.* 1994a, 1994b). While Hanemann's model introduces the theoretically desirable connection between income and demand, it has been found to be wanting in other respects. It may produce unexpected negative benefit measures (Bowker and Stoll 1988) and positively sloped demand curves (Boyle and Bishop 1988), and is inferior to log-logistic models on the statistical fit criterion. These results lead Bowker and Stoll (1988, p. 379) to call for 'better specifications of the utility function', a call recently echoed by Kanninen (1995, pp. 120–1), who states that 'the researcher should experiment with alternative functional forms for the WTP distribution. The existence of thick upper tails suggests the use of an asymmetric distribution which might be less sensitive to bids in the tails.'

There are several issues in choice of a parametric distribution for the representative demand function, including: goodness-of-fit, theoretical consistency, and mathematical tractability. The difficulty of reconciling the first two aims is illustrated by cases in which the log-logistic distribution outperforms the logistic distribution on goodness-of-fit criteria, but yields infinite mean consumer surplus estimates.[3] The principal approaches to dealing with this problem include Winsorizing[4] (Duffield and Patterson 1991), and truncation[5] at some 'reasonable' upper bid limit, sometimes

[2] Known, respectively, as logistic and log-logistic models.

[3] Log-logistic models do not always yield infinite mean consumer surplus, but can do so with some parameter estimates (Hanemann 1984). It may be more correct to claim that mean consumer surplus is undefined, rather than infinite. The log-logistic function is asymptotic to the bid axis, but in some cases the area under the function converges to a finite limit. In cases where the log-logistic function is 'flatter', such convergence does not occur, and in this sense the area under the curve (mean WTP) can be interpreted as infinite. Whichever interpretation is adopted, the result is problematic.

[4] Duffield and Patterson (1991) explain Winsorizing as 'assign a value of T to all WTP values above T before computing the mean'. This approach is also referred to as censoring at T.

[5] Truncation entails dropping from analysis all bids outside some (arbitrary) bid value range.

accompanied by normalisation[6] as an attempt to restore other aspects of theoretical consistency (Boyle, Welsh and Bishop 1988). It is also possible to choose distributions with alternative upper tail shapes, as suggested by Kanninen, or to manipulate distributions to force desirable tail characteristics (Ready and Hu 1995). Concern about the lower tail shape has given rise to spike models, which are designed to accommodate discontinuities that arise because large segments of the population have no interest in the item being valued (i.e. they are 'not in the market'), or are losers if the proposed change proceeds (Kriström 1997).

The major logistic distribution-based models are used as approximations to theory-consistent functional forms (Bowker and Stoll 1988; Duffield and Patterson 1991; Johansson, Kriström and Maler 1989; Park, Loomis and Creel 1991; Sellar, Chavas and Stoll 1986). Other distributions may work as well as, or better than, logistic distribution-based models in terms of one or more of the goodness-of-fit, theoretical consistency, and mathematical tractability criteria. This article provides an initial investigation of parameterisation effects. It tests variants of the logistic and log-logistic distributions to identify their ability to conform to prior expectations. It also investigates the implications of several variants of the Weibull/exponential family of probability distributions. All models are fitted to an existing dichotomous choice data set in order to highlight their implications.

2. Survival functions

The dichotomous choice approach to contingent valuation determines the probability of a representative consumer being willing to pay some nominated amount to ensure that a given environmental change occurs, or is avoided. Alternatively, the probability of willingness to accept some given amount of compensation may be estimated. As the nominated money amount is increased, the proportion of respondents willing to pay for the change is expected to decrease. The 'survival function' depicts how likelihood of WTP decreases as bids increase. The logit and probit models are consistent with the observation that probabilities are constrained to be in the (inclusive) range between zero and one.[7] However, neither of these distributions allows WTP to be positive for everyone at some positive bid level nor to decline to zero at non-infinite bids. Both these outcomes are commonly observed and have been proposed elsewhere as minimum criteria for a valid model of WTP (Haab and McConnell 1997, 1998).

[6] Boyle *et al.* (1988) suggested a normalisation process to ensure that probability of WTP falls to zero at the upper truncation point of the bid range.

[7] Although neither of these models allows probabilities to equal either zero or one.

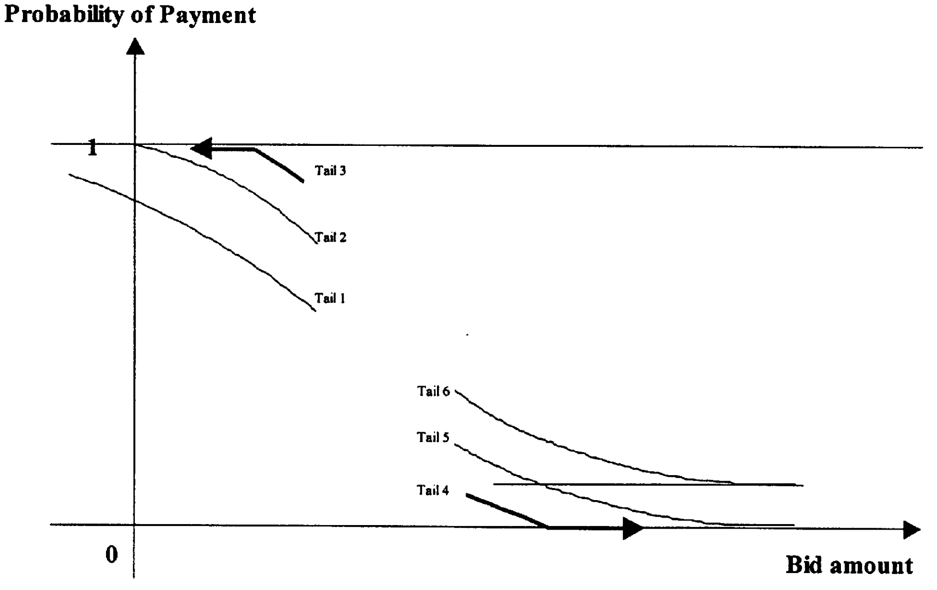

Figure 1 Possible tail shapes for survival functions

Minimal desirable conditions that should be satisfied by all survival functions include finite WTP and decreasing likelihood of payment as bid increases. Additionally, tail shapes may need to vary to accommodate differences in response patterns.

Consider the lower (left) tail of the survival function.[8] The survival function may take one of three alternative tail forms at this point:

Tail 1: intersect the probability axis at a probability less than one.
Tail 2: intersect the probability axis at a probability equal to one.
Tail 3: probability reaches one at a bid amount greater than zero.

These alternative shapes are presented in figure 1. The logistic model (among others) represents Tail 1, while the log-logistic model represents Tail 2.

The survival function should exhibit Tail 1 for those cases in which some people are not willing to 'purchase' the non-market good at the current price.[9] Measurement of existence values provides an example. If some people truly do not care about, or have preferences for the destruction of, the resource, then measurement of WTP for the population as a whole will yield

[8] The left tail is that part of the distribution about the point where bid is zero.

[9] Current price might be, for example, travel and/or time costs for recreational resource use.

responses of this type (Kriström 1997). In the words of Haab and McConnell (1998, p. 219), 'In the case of most public goods, the good being valued can simply be ignored if it does not provide an increase in utility.' Consequently, there may be no individuals with negative WTP, but a non-zero mass of people WTP nothing.

Tail 2 is expected to occur in situations where user populations alone are surveyed, e.g. on-site surveys of recreationists. Most recreationists will have positive WTP (above current costs) for their recreation activity, but the marginal recreationist would quit the activity if price were increased marginally. Tail 1 type data may be transformed to Tail 2 type data by excluding all of those people who are not willing to pay marginally more than the present price. This requires appropriate screening questions within the contingent valuation survey.

The introduction of either non-price rationing or uncertainty is sufficient to justify Tail 3. Where access to a resource is limited, say, by membership of a closed club or by government policy, marginal WTP could be greater than zero. Another example is provided by experiences that are superior to expectations. In such cases, WTP of every resource user may exceed zero (i.e. the *ex ante* marginal user is not ambivalent, *ex post*, about the decision to utilise the resource); however, this may only be a transitory phase while expectations adjust.[10] Tail type 3 may also be observed within particular populations because of self-selection, e.g. environmental activists' WTP to preserve tropical rain forests.

The policy implications of Tail 1 have been the stimulus for some debate about whether the mean of all WTP values, the mean of only positive WTP values, or the median provides the most appropriate aggregate benefit measure (Hanemann 1984, 1989; Johansson *et al.* 1989). Kriström (1997) broadens this argument by introducing the prospect of two types of spike model, one in which some people are not in the market (the good does not enter their utility function), and another in which some people are winners, some are losers and some are ambivalent. Where people are not in the market, inclusion of negative bids derived from parametric model extrapolation causes mean WTP (and possibly the median) to be underestimated. The second spike model, which includes winners and losers, suggests a discontinuity at zero that may require a different model specification on each side of the discontinuity.

[10] On the other hand, expectations may exceed the benefits actually delivered by a non-market resource, a situation that may be exacerbated by marketing strategies. If feedback loops are poor or non-existent (e.g. when recreational activities are purchased by international travellers and there is little return custom), unrealised expectations may be an ongoing phenomenon. In the presence of unrealised expectations Alternative 1 may be observed, whereas Alternative 2 is expected in cases where people are perfectly informed.

There are also three alternatives for the shape of the low probability end (upper tail) of the survival function:

Tail 4: reaches the bid axis at a finite bid.
Tail 5: asymptotic to the bid axis.
Tail 6: asymptotic to a probability greater than zero.

The logistic and log-logistic models are both examples of Tail 5. They imply that some people are willing to pay infinite money amounts, which contravenes prior expectations for the types of goods usually valued using these functions. This is termed the 'fat tail problem' by Boyle *et al.* (1988). The logistic model presents less of a problem because, while it allows some people to have infinite WTP, mean consumer's surplus is finite. This is not the case for some coefficient values of the log-logistic model, for which the mean is non-finite (Hanemann 1984). Tail 4 is consistent with the desirable characteristics of survival functions and always provides finite mean consumer's surplus. Tail 6 is undesirable because it always yields infinite mean WTP. Ready and Hu's (1995) pinched log-logistic model is one approach for transforming Tail 5 to Tail 4.

With three possibilities for each of the two tails it is possible to construct nine different types of survival function. However, three of those nine (incorporating Tail 6) are rejected because of infinite consumer surplus estimates. A further three (incorporating Tail 5) have the undesirable characteristic of allowing individuals to have infinite WTP, but may provide useful approximations in cases where there is a finite limit of integration.

3. Survival function variants

A host of distributions, in addition to the already used normal, logistic, log-logistic and Weibull distributions, is available to the dichotomous choice analyst, but many of these can be ruled out on theoretical or pragmatic grounds. For example, while the Cauchy and Type I Extreme Value[11] functions often provide very good fits to data and are extremely tractable mathematically, they both yield infinite mean WTP (Johnson and Kotz 1971).[12] For those cases in which a measure of mean WTP is required, say, as an input to cost-benefit analysis, these distributions will be unhelpful. However, they may be beneficial in cases where median WTP is sought, say,

---

[11] Sometimes also referred to as 'log-Weibull' and 'double exponential'.

[12] The probabilities of YES responses (survival rates) for these functions are:

Cauchy $P_{YES} = 0.5 - \tan^{-1}(A + BX)/\pi$

Type I Extreme Value $P_{YES} = \exp(-\exp(A + BX))$

to predict referendum results or to assess citizen acceptance of a proposal. These distributions are not examined further here.

As an alternative to searching across the range of cumulative density functions that might potentially allow the desired flexibility in survival function tail shapes, the approach adopted here is to explore the benefits of variations on survival functions that are already in wide use.[13] Following the approach used by Boyle *et al.* (1988) and Ready and Hu (1995), additional parameters are added to commonly used distributions to allow them to comply with expected tail shapes and to better fit the data. The survival function variants are first introduced and some key descriptors are defined, then the range of functions is applied to data to reveal the dependence of the descriptors on survival function variant.

Scaling provides flexibility in the lower tail. Scaling has the effect of multiplying predicted probabilities from the standard form of the survival function by a constant factor. Let $g(\mathrm{X})$ be the standard survival function, which predicts probability of WTP, $f(\mathrm{X})$, as a function of bid (X), i.e. $f(\mathrm{X}) = g(\mathrm{X})$. The scaled survival function has the form $f(\mathrm{X}) = \alpha.g(\mathrm{X})$, where α is some constant. This simple change allows the lower tail to take any of the 3 potential forms, but does not affect the general shape of the upper tail.[14]

Compared to the standard form of the survival function, a shifted function increases all probabilities by some fixed (possibly negative amount). The shifted model has the form $f(\mathrm{X}) = g(\mathrm{X}) + \beta$, where β is some constant. Addition of a shift parameter allows some flexibility in the upper tail, but only at the cost of a shift in the lower tail (and vice versa). For example, a shifted log-logistic model that reaches the bid axis *must* have probability of WTP of less than unity at a zero bid. Shifted models are therefore rather clumsy and may consequently not substantially increase explanatory power over the standard forms.

The changes in survival function shape because of scaling and shifting are illustrated in figures 2 and 3, respectively. Figure 2 shows scaling has its largest effect at low bids, and has a small effect at high bids. Scale factors greater than one increase probabilities of bid acceptance, all else being unchanged. However, the addition of a scale parameter will normally affect all estimated parameters and will consequently change the overall shape of the fitted survival function. Figure 3 illustrates the uniform impact of the

[13] These are the logistic, log-logistic, Weibull and exponential distributions. The normal distribution is not included because of its similarity to the logistic distribution. McFadden (1994) describes and Cooper (1993) utilises the gamma distribution. The beta model too has been introduced recently (Haab and McConnell 1998). Neither the beta nor the gamma distribution is analysed here.

[14] Scaling is implicit in the distributions described in McFadden (1994).

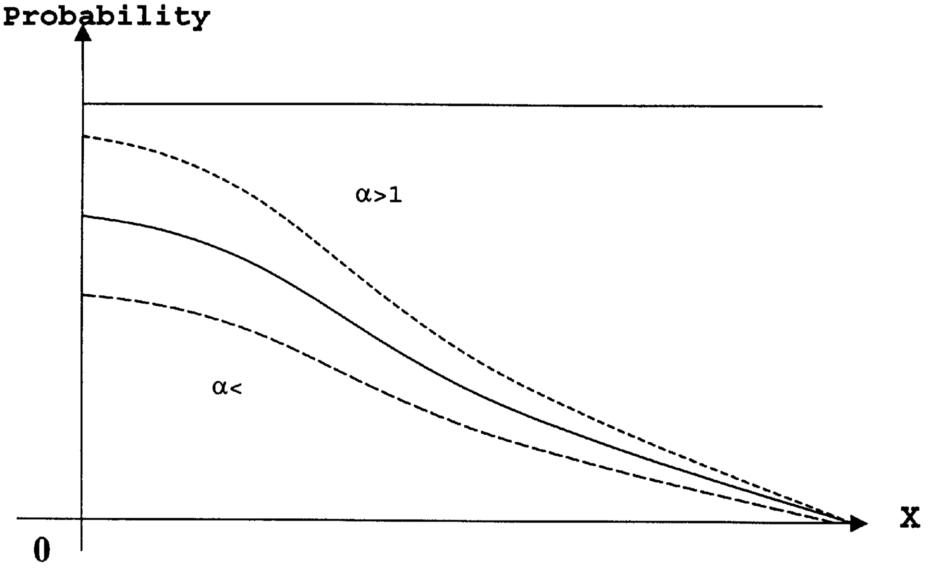

Figure 2 Scaled survival functions

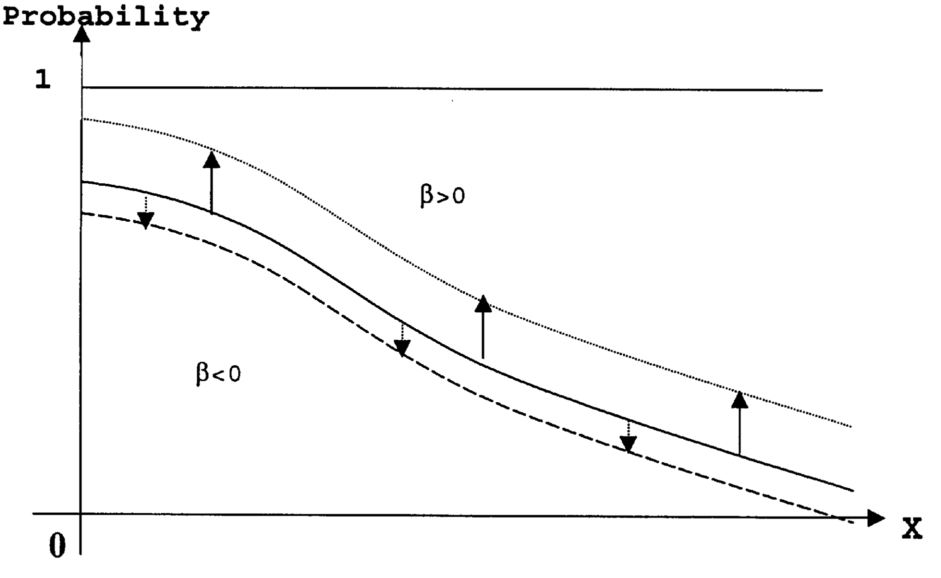

Figure 3 Shifted survival functions

shift parameter across the whole survival function. Shift factors greater than zero can be envisaged as 'sliding' the survival function up the y-axis. Again, the overall shape of the fitted function is able to change with the addition of a shift parameter.

One approach to obtaining the upper tail benefits of shifting as well as the lower tail benefits of scaling is to combine both into a single specification of the survival function. The generalised model incorporates both scale and shift parameters and has the form $f(\mathrm{X}) = \alpha.g(\mathrm{X}) + \beta$, where α and β are constants. All tail combinations are possible and the generalised model produces the standard, scaled and shifted models as special case outcomes.

The variants of the logistic survival function illustrate how the additional parameters enter the survival functions. These are:

Standard logistic survival function	$f(\mathrm{X}) = [1 + \mathrm{e}^{-(\mathrm{A+BX})}]^{-1}$
Scaled logistic survival function	$f(\mathrm{X}) = \alpha.[1 + \mathrm{e}^{-(\mathrm{A+BX})}]^{-1}$
Shifted logistic survival function	$f(\mathrm{X}) = [1 + \mathrm{e}^{-(\mathrm{A+BX})}]^{-1} + \beta$
Generalised logistic survival function	$f(\mathrm{X}) = \alpha.[1 + \mathrm{e}^{-(\mathrm{A+BX})}]^{-1} + \beta$

In each case A is a constant, B is the coefficient on the money bid, α is the scale parameter and β is the shift parameter.

The implications of these changes in model specification may be quite profound. For example, the upper tails of the shifted and generalised functions for the commonly used survival functions are asymptotic to probability equal to β, so mean consumer's surplus is infinite whenever $\beta > 0$, and is finite whenever $\beta < 0$.

A further potential variation on currently used survival functions involves a transformation of the upper tail, as illustrated by the 'pinched log-logistic' form utilised by Ready and Hu (1995), $f(\mathrm{X}) = (1 - \mathrm{X/T})g(\mathrm{X})$, where T is the (endogenous) point at which the upper tail reaches the bid axis and $g(\mathrm{X})$ is the standard log-logistic survival function. This model retains the constraint that probability of paying a zero bid is one, but ensures that probability becomes zero at the finite bid level, T. In contrast, Kriström's (1997) spike model transforms the lower tail while retaining the original upper tail shape.

The variants do not constrain $f(\mathrm{X})$ to fall within the range (0, 1). Let $\mathrm{P_{YES}(X)}$ be the probability of WTP money bid amount X, i.e. the survival function. Then:

$$\mathrm{P_{YES}(X)} = 1 \qquad \text{if } f(\mathrm{X}) \geq 1 \tag{1}$$

$$\mathrm{P_{YES}(X)} = f(\mathrm{X}) \qquad \text{if } 0 < f(\mathrm{X}) < 1 \tag{2}$$

$$\mathrm{P_{YES}(X)} = 0 \qquad \text{if } f(\mathrm{X}) \leq 0 \tag{3}$$

4. Derivation of mean WTP

For Tail 1, different derivations of the mean apply, depending upon whether negative WTP is permitted or not. The following derivations assume that negative WTP is not permitted. To evaluate mean WTP for scaled, shifted and generalised models it is necessary to define X^L and X^U, the bids at which probabilities of WTP are 1 and 0, respectively.

$$f(X^L) = 1, X^L \geq 0, \text{and}$$

$$f(X^U) = 0$$

To illustrate, X^L and X^U are derived for the scaled logistic model by solving for X in the following equations.

$$X^L: \quad C.[1 + e^{-(A+BX)}]^{-1} = 1 \quad \Rightarrow \quad X^L = -[A + \ln(C-1)]/B \quad \{\text{for } C > 1\}$$

$$X^U: \quad C.[1 + e^{-(A+BX)}]^{-1} = 0 \quad \Rightarrow \quad X^U = \infty$$

Mean WTP is infinite whenever X^U is not defined (Tail 6). This condition occurs only for shifted and generalised models. Recalling that X^L is non-negative, mean WTP is the area under the survival function:

$$\text{Mean} = X^L + \int_{X^L}^{X^U} f(X)dX$$

It is often easier to evaluate the equivalent:

$$\text{Mean} = X^L + \int_0^{X^U} f(X)dX - \int_0^{X^L} f(X)dX$$

The integrals may need to be evaluated numerically. Simple analytical solutions for median WTP exist for all model variants.[15]

5. Empirical test of impacts

The implications of alternative functional forms are tested by their application to data from a household survey used to determine benefits of water quality improvement in New Zealand's Waimakariri River (Sheppard *et al.* 1993). The Waimakariri River runs to the sea on the outskirts of Christchurch City. It receives substantial recreational use from residents of Christchurch and the surrounding region. It also receives several major pollutant discharges that reduce the recreational and aesthetic values of the

[15] Survival function definitions and analytical solutions for X^L, X^U, Median WTP and Mean WTP are available from the author on request.

lower river reaches. The study used off-site surveys to value an improvement in water quality. Because an off-site survey would have captured some respondents who had no interest in the issue,[16] Tail 1 is expected. The data confirm this expectation, with 16.3 per cent of valid responses indicating their non-willingness to pay the nominal sum of $1 for the proposed water quality improvement. Shift parameters less than unity should be found in the log-logistic, Weibull and exponential distribution-based models which ordinarily assume 100 per cent WTP when the bid is zero.

While the survey did not contain questions directly analogous to Kriström's questioning format, it is assumed that the contingent market validity test applied to verify validity of responses indicating non-WTP $1 is a suitable proxy for application of the simple 'not in the market' spike model.

After validation of willingness to enter the contingent market, survey participants responded to a single dichotomous choice question that sought their response to a referendum. The two available responses were (1) the status quo, and (2) payment of additional annual regional council rates to treat discharges into the river, thereby raising the water quality from its existing suitability only for fishing and boating to a level that would be safe for swimming. A postal survey was administered that obtained 1 161 responses from 2 628 delivered questionnaires (44 per cent). Follow-up telephone interviews of non-respondents indicated that survey responses were biased towards river visitors. After deletion of item non-responses and protest responses, 824 cases remained to be analysed. Response rates did not vary significantly across the range of bids. Data are reported in table 1.

Pre-tests indicated the likely range of bids, and the bid amounts subsequently included in the survey were chosen in the expectation that the higher amounts would exceed WTP of nearly every individual in the population. This expectation proved to be false, with an unexpected number of positive responses to the high bids. These data therefore represent the classic fat tails situation, in which the data are not optimal, but are nevertheless the basis for some policy evaluation task that requires measurement of the area under the survival function.

Parametric models have been fitted using maximum likelihood estimation. Significance of improvements in fit are tested using likelihood ratio tests (Cramer 1986). If the standard models that are commonly used are poor approximations to the underlying distributions of WTP, it is expected that at least some of the descriptors (mean, median, probability of WTP zero, X^U) will change significantly with the inclusion of additional parameters.

[16] The lower Waimakariri River had been visited by 47 per cent of survey respondents (Sheppard *et al.* 1993).

Table 1 Data

Dollar amount	Number WTP	Number not WTP	Proportion WTP	Number of responses at this bid amount
2	25	2	.926	27
7	20	6	.769	26
12	20	5	.800	25
17	19	6	.760	25
22	12	14	.462	26
27	13	8	.619	21
32	16	12	.571	28
37	15	14	.517	29
42	6	16	.273	22
47	15	10	.600	25
57	21	35	.375	56
67	18	26	.409	44
77	29	32	.475	61
87	26	30	.464	56
97	16	38	.281	54
117	10	20	.333	30
137	8	16	.333	24
157	11	19	.367	30
177	12	23	.343	35
197	5	19	.208	24
217	6	33	.154	39
237	7	25	.219	32
257	3	29	.094	32
277	8	17	.320	25
297	2	26	.071	28
Total	343	481		824

6. Results

Descriptor variables and their bootstrapped confidence intervals are reported in table 2 and likelihood ratio tests of additional parameters are presented in table 3. Point estimates of the mean (defined here only over the positive bid range) vary dramatically with functional form, ranging from \$94 to infinity. Excluding models that predict infinite means, the highest point estimate of the mean is \$308. Differences in non-infinite means are statistically significant in some cases. Upper 95 per cent confidence interval bounds for the relatively poorly fitting logistic, spiked logistic, exponential and shifted exponential models are \$130, \$103, \$108 and \$120, respectively. Comparison with lower 95 per cent confidence interval bounds for shifted log-logistic, Weibull and scaled Weibull models (\$121, \$128, \$169) indicates a number of significant differences. For comparative purposes, Kriström's (1990) non-parametric (PAVA) estimate of the mean for these data, using an

Table 2 Results

Model	Log-likelihood	Median WTP ($)	Mean WTP ($)	X^U ($)	Probability of WTP zero
Constant	559.5				0.42
Logistic	509.6	64 (45,80)	111 (97,130)	$\underline{\infty}$	0.64 (0.59,0.70)
Spiked logistic	537.7	81 (73,91)	94 (86,103)	$\underline{\infty}$	0.80 (0.78,0.82)
Scaled logistic #	505.3	61	116	$\underline{\infty}$	0.73
Shifted logistic	504.3	49 (34,71)	∞ (∞,∞)	nd (nd,nd)	0.79 (0.65,0.97)
Generalised logistic #	501.7	50	∞	nd	0.86
Log-logistic	498.7	44 (34,54)	∞ (∞,∞)	$\underline{\infty}$	$\underline{1.00}$
Pinched log-logistic	498.4	46 (36,56)	134 (100,9431)	820 (399,1.7×10^8)	$\underline{1.00}$
Scaled log-logistic	498.7	44 (27,56)	∞ (2715, ∞)	$\underline{\infty}$	1.00 (0.81,1.00)
Shifted log-logistic	498.7	44 (33,58)	308 (121, ∞)	12310 (776,nd)	0.99 (0.83,1.00)
Generalised log-logistic #	498.0	45	123	644	1.00
Weibull	498.7	48 (37,59)	171 (128,292)	$\underline{\infty}$	$\underline{1.00}$
Scaled Weibull	498.3	45 (34,60)	215 (169,1318)	$\underline{\infty}$	1.00 (0.86,1.00)
Shifted Weibull	498.5	46 (35,60)	∞ (605,∞)	nd (605,nd)	1.00 (0.87,1.00)
Generalised Weibull #	498.0	45	123	644	1.00
Exponential	531.1	69 (64,75)	100 (92,108)	$\underline{\infty}$	$\underline{1.00}$
Scaled Exponential	505.2	60 (47,72)	116 (101,138)	$\underline{\infty}$	0.73 (0.65,0.82)
Shifted Exponential	508.3	66 (48,79)	105 (95,120)	380 (342,441)	0.68 (0.61,0.76)
Generalised Exponential	501.7	50 (35,64)	∞ (108,∞)	nd (577,nd)	0.86 (0.72,0.99)

Notes:
Underlined entries (e.g. $\underline{1.00}$) take their values by definition.
nd Not defined (e.g. requires logarithm of a negative number, or the bid function is asymptotic to a positive probability).
Confidence intervals have not been derived because of convergence problems in these models.

Table 3 Chi squared tests of significance of additional parameters

Variant	Degrees of freedom	Logistic	Log-logistic	Weibull	Exponential
Standard vs pinched	1		0.62		
Standard vs scaled	1	8.74**	0.01	0.83	51.87**
Standard vs shifted	1	10.72**	0.04	0.42	45.54**
Standard vs generalised	2	15.78**	1.36	1.44	58.71**
Pinched vs generalised	1		0.75		
Scaled vs generalised	1	7.05**	1.36	0.61	6.84**
Shifted vs generalised	1	5.06*	1.33	1.02	13.17**

Notes:
* Significant @ a = 0.05
** Significant @ a = 0.01

Critical values:

	a = .05	a = .01
1 dof	3.84	6.63
2 dof	5.99	9.21
3 dof	7.81	11.34

upper bid limit of $350 obtained by linear extrapolation, is $103. It is notable, however, that the lower bound measures of mean WTP (for non-infinite cases) fall within a narrow range, apart from the shifted Weibull and scaled log-logistic models.

Median estimates show consistency, differences are generally not significant, with the log-logistic and Weibull-based models showing remarkable consistency. Medians for the logistic and exponential models tend to be higher than for the other models, but this difference is only significant for the poorly fitting one-parameter, standard exponential model and the even poorer fitting spiked logistic model. The symmetric logistic distribution does not reflect the asymmetric nature of the data, which may have constrained this model from reliably estimating the median.[17]

Addition of parameters had little impact on goodness-of-fit of the log-logistic and Weibull forms, but offered significant improvements for the poorer fitting logistic and exponential forms. The generalised models performed poorly. The generalised exponential form approached the explanatory power of the log-logistic and Weibull families of functions. However, the generalised exponential model also yielded undesirable Tail 6. The other generalised models all exhibited convergence problems and, while it was possible to derive (unreliable) point estimates of parameters, it was not possible to derive confidence intervals. Similar difficulties were en-

[17] I am indebted to an anonymous reviewer for this point.

countered with the scaled logistic model. The shifted logistic and shifted Weibull models produced Tail 6, while the shifted log-logistic and shifted exponential models produced a transition from Tail 5 to Tail 4. In obtaining this desirable outcome the shifted exponential model has over-corrected the left tail so that the confidence interval for probability of WTP at zero bid does not overlap that observed in the data.

A priori expectations were for probability to be less than unity when the bid amount is zero. This expectation is borne out by the logistic and exponential-based models, with differences from unity always being significant. The standard logistic model predicts a somewhat lower zero-bid probability than other models, significantly so in some cases. Probabilities are not significantly different from unity for the log-logistic and Weibull-based models, although lower bound values are all less than one and are close to the observed value. The upper bound on X^L over all models was \$2. Clearly, Tail 3 does not apply to these data. However it is not possible to distinguish statistically between Tails 1 and 2, both being supported by some model specifications. It is notable that the spiked logistic model, although coming very close to correctly fitting the left tail, offers an extremely poorly fitting survival function.

7. Discussion

The fat tails problem has been known for some time and a variety of approaches, principally based on truncation, have been used to address it. Recognition of the problem is implicit acceptance that the parametric models being fitted to the data do not well represent expected and/or observed behaviour. At least three potential causes arise: (1) the array of bids presented to survey participants is inadequate to define the tails of the distribution; (2) incorrect parameterisation of fitted models; or (3) the data are poorly behaved because of strategic or other behavioural responses.

Better anchoring the tail of the distribution would seem to imply the desirability of placing more bids in the upper tail. However, this solution is not always possible in applied settings where time and budget constraints often force reliance on existing data. Even with extensive survey pre-testing it is not uncommon to find surprises in the data, typically an unexpected high frequency of yes responses to high bids. In many instances it is simply not possible for practical reasons to utilise the information contained in these data to undertake further sampling to address data concerns, either by expanding the sample size to reduce the probability of ‘flyers’ or to alter the bid distribution. However, measures of value are often still required to help resolve some resource management issue and the existing data may provide the only source of such estimates, forcing reliance on less than optimal data.

Further sampling to develop an optimal bid distribution may be of little value anyway. Several authors have tackled the problem of identification of the optimal bid distribution for dichotomous choice contingent valuation (Duffield and Patterson 1991; Cooper and Loomis 1992, 1993; Cooper 1993; Kanninen and Kriström 1993; Alberini 1995a; Kanninen 1995). These studies generally make strong assumptions about the form of the underlying bid distribution as the basis for their analyses. As Kanninen and Kriström (1993, p. 201, emphasis added) indicate 'it is not necessary, *provided the distributional assumption is correct*, to cover the entire range of WTP values'. However, it is apparent that in applied settings the underlying survival function is unknown *a priori*. Alberini (1995b) used simulated data to test the ability to differentiate between correct and incorrect distributions. She found that extending the vector of bids did not increase the power to identify a mis-specified distribution except in some instances when sample size was very large. Alberini's test is not conclusive, however, as she used an optimal sampling routine based on each of the assumed distributions[18] so that the sample differed between data sets, unlike most applied cases in which a number of distributions are fitted to a single data set. Where the survival function is unknown *a priori*, neither theory nor practice provide clear guidance on optimal dichotomous choice bid distributions.

The possibility of incorrect survival function parameterisation suggests testing alternative distributions. This article has done so by using four common distributions, and by increasing their flexibility. The simple changes made to commonly used distributions are additions to a process of distributional modification. Results have been mixed. Variants of the log-logistic and Weibull distributions offer negligible advantage over other forms in terms of goodness-of-fit. Adding parameters only offered significant improvements for the poorer fitting logistic and exponential forms. Standard log-logistic distributions commonly fail to yield finite mean WTP. In this case, scaling failed to address that problem, but shifted and generalised forms of the log-logistic distribution were partially successful. Both means had finite expected values, although the upper bound was not defined for the shifted model, and no confidence interval could be defined for the generalised model.

Choice of the best model cannot be made on statistical grounds as there is very little, if any, difference in goodness-of-fit. Similarity of goodness-of-fit coupled with non-overlapping estimates of the mean between distributional forms raises the issue of which estimate of the mean should be used for cost-benefit analysis. The fat tail problem remains in somewhat different form,

[18] The two distributions tested were log-normal and Weibull.

there exists a variety of distributions which satisfy the non-infinite WTP criterion, but these provide somewhat diverse benefit estimates — the ratio of highest to lowest non-infinite point estimates derived here being about 3. The choice between competing functional forms can be resolved statistically for the bid range for which data exist, but the behaviour of survival functions outside that range cannot be judged. Further search for distributions which resolve the problem is therefore likely to be unsuccessful in that the 'correct' distribution could already be within the set of models tested here, but the analyst has no way of identifying it. In other words, it is not possible to estimate our way out of the fat tail problem.

Several authors note the potential role of yea-saying in the fat-tail problem (Alberini 1995b; Kanninen 1995). Yea-saying, nay-saying and other non-truthful response behaviours[19] such as strategic responses, random responses, faulty coding or any other actions resulting in 'non-truthful' data also possess the potential to induce the fat tail problem and should not be overlooked. Any of these actions could effectively result in Tail 6, even though true preferences embody Tail 4. If the inclusion of extended bid ranges creates more incentives for non-truthful responses or does not remove the reasons for this type of behaviour, then extension of the bid range will not remove the fat tail problem, it may simply bias results. For example, if very high bids are included, these may seem quite implausible to some respondents who may then respond flippantly or strategically. While non-truthful responses are consistent with the unexpected frequency of yes responses to high bids in the Waimakariri River study, the high variance in responses across the entire bid range suggests that this potential explanation should not be adopted unquestioningly. It should also be noted that neither distributional variety nor extended bid ranges adequately address issues arising from fat tail problems that result because of behavioural response issues.

Estimates of benefits from water quality improvement in the Lower Waimakariri River varied by a factor of three over the range of theoretically acceptable distributions fitted to the data. The results presented here bring into question the robustness of non-market benefit estimates that have been derived using only one or two distributions. It is readily apparent that the variability of benefit measures requires caution in the use of such values to inform policy decisions and, in the short term at least, a need for greater sensitivity analysis on the part of dichotomous choice analysts. This study has investigated a very small range of the distributions upon which that sensitivity analysis might be based. There may be other distributions that

[19] Whether intentional or not.

would fit the data much better, removing many of the concerns raised here, however, further research in this direction must eliminate yea-saying and other behavioural causes of the fat tail problem in order to make progress.

The consistency of lower bound mean WTP estimates suggests that use of lower bounds to support a change yielding non-market benefits (e.g., improving water quality in the Lower Waimakariri River) is justified. The volatility of upper bounds implies the converse does not hold.

It is worth noting the consistency of medians estimated across functional forms. There were only two outliers, the spiked logistic model and the extremely simple standard exponential model. The single parameter exponential model cannot be expected to fit the data well and is unlikely to be utilised. However, the spiked logistic model has been promoted as a desirable alternative to other commonly used models. Inability of the symmetric logistic model both to conform to the spike and to fit the remainder of the data has resulted in the highest point estimate of the median (and the lowest estimate of the mean). This result signals the need for caution in utilisation of measures derived using the spiked logistic model. Notwithstanding the discrepancies already noted, the median is extremely robust to assumptions about functional form in this case.

It has not been possible to identify the relative abilities of the new distributional variants to estimate true mean WTP because that quantity is unknown under the case study approach used here. The use of synthetic data of known form and mean offers an alternative approach for testing the potential advantages of these new distributions. It also allows for elimination of behavioural response biases. The resolution of fat tail problems is dependent upon their cause, implying the need for better understanding of dichotomous choice behaviours. Even when the data represent 'true' responses, the fat tail problem may remain. These results suggest the need for a great deal of caution in reliance on dichotomous choice results where it has not been shown that mean WTP is insensitive to distributional assumption. It does not appear possible to remove this problem by adding flexibility to commonly used distributions to better accommodate both economic theory and ability to conform with empirical data, although the potential benefits of extending the search to other distributions cannot be ruled out.

The opportunity for 'art' in the design and interpretation of dichotomous choice contingent valuation studies remains. Analysts need to be aware of the nature of their data and should choose parameterisations accordingly. For example, shifted models are incapable of yielding survival functions that simultaneously have positive lower bounds and non-infinite upper bounds. Approaches to dealing with problems perceived in one tail of the survival function have implications for the other tail and the fit to the rest of the data.

Awareness of these implications indicates the need for close scrutiny of assumptions about the form of the data and choice of a survival function that conforms with these assumptions. Questionnaire design that explicitly identifies tail shape, as with the spike models, is likely to be extremely useful in this regard. Median WTP appears relatively immune to changes in functional form, but close scrutiny and increased sensitivity analysis are urged for applications reliant upon dichotomous choice estimates of mean WTP.

References

Alberini, A. 1995a, 'Optimal designs for discrete choice contingent valuation surveys: single-bound, double-bound, and bivariate models', *Journal of Environmental Economics and Management,* vol. 28, pp. 287–306.

Alberini, A. 1995b, 'Testing willingness-to-pay models of discrete choice contingent valuation survey data', *Land Economics,* vol. 71, pp. 83–95.

Bowker, J.M. and Stoll, J.R. 1988, 'Dichotomous choice in resource valuation', *American Journal of Agricultural Economics,* vol. 70, pp. 372–81.

Boyle, K.J. and Bishop, R.C. 1988, 'Welfare measurements using contingent valuation: a comparison of techniques', *American Journal of Agricultural Economics,* vol. 70, pp. 20–8.

Boyle, K.J., Welsh, M.P. and Bishop, R.C. 1988, 'Validation of empirical measures of welfare change: comment', *Land Economics,* vol. 64, pp. 94–8.

Carson, R.T., Mitchell, R.C., Hanemann, W.M., Kopp, R.J., Presser, S. and Ruud, P.A. 1994a, *Contingent Valuation and Lost Passive Use: Damages from the Exxon Valdez.* Resources for the Future Discussion Paper, 94–18.

Carson, R.T., Wilks, L. and Imber, D. 1994b, 'Valuing the preservation of Australia's Kakadu Conservation Zone', *Oxford Economic Papers,* vol. 46, (Supplementary issue), pp. 727–49.

Cooper, J. and Loomis, J. 1992, 'Sensitivity of willingness-to-pay estimates to bid design in dichotomous choice contingent valuation models', *Land Economics,* vol. 68, pp. 211–24.

Cooper, J. and Loomis, J. 1993, 'Sensitivity of willingness-to-pay estimates to bid design in dichotomous choice contingent valuation models: reply', *Land Economics*, vol. 69, pp. 203–8.

Cooper, J.C. 1993, 'Optimal bid selection for dichotomous choice contingent valuation surveys', *Journal of Environmental Economics and Management,* vol. 24, pp. 25–40.

Cramer, J.S. 1986, *Econometric Applications of Maximum Likelihood Methods,* Cambridge University Press, Cambridge.

Creel, M. and Loomis, J. 1997 'Semi-nonparametric distribution-free dichotomous choice contingent valuation', *Journal of Environmental Economics and Management,* vol. 32, pp. 341–58.

Duffield, J.W. and Patterson, D.A. 1991, 'Inference and optimal design for a welfare measure in dichotomous choice contingent valuation', *Land Economics,* vol. 67, pp. 225–39.

Haab, T.C. and McConnell, K.E. 1997, 'Referendum models and negative willingness to pay: alternative solutions', *Journal of Environmental Economics and Management,* vol. 32, pp. 251–70.

Haab, T.C. and McConnell, K.E. 1998, 'Referendum models and economic values: theoretical, intuitive and practical bounds on willingness to pay', *Land Economics,* vol. 74, pp. 216–29.

Hanemann, W.M. 1984, 'Welfare evaluations in contingent valuation experiments with discrete responses', *American Journal of Agricultural Economics,* vol. 66, pp. 332–41.

Hanemann, W.M. 1989, 'Welfare evaluations in contingent valuation experiments with discrete response data: reply', *American Journal of Agricultural Economics,* vol. 71, pp. 1057–61.

Johansson, P.-O., Kriström, B. and Maler, K.G. 1989, 'Welfare evaluations in contingent valuation experiments with discrete response data: comment', *American Journal of Agricultural Economics,* vol. 71, pp. 1054–56.

Johnson, N.L. and Kotz, S. 1971, *Continuous Univariate Distributions: 2,* Wiley and Sons, New York.

Kanninen, B.J. 1995, 'Bias in discrete response contingent valuation', *Journal of Environmental Economics and Management,* vol. 28, pp. 114–25.

Kanninen, B.J. and Kriström, B. 1993, 'Sensitivity of willingness-to-pay estimates to bid design in dichotomous choice valuation models: comment', *Land Economics,* vol. 69, pp. 199–202.

Kriström, B. 1990, 'A non-parametric approach to the estimation of welfare measures in discrete response valuation studies', *Land Economics*, vol. 66, pp. 135–9.

Kriström, B. 1997, 'Spike models in contingent valuation', *American Journal of Agricultural Economics*, vol. 79, pp. 1013–23.

McFadden, D. 1994, 'Contingent valuation and social choice', *American Journal of Agricultural Economics,* vol. 76, pp. 689–708

Park, T., Loomis, J.B. and Creel, M. 1991, 'Confidence intervals for evaluating benefits estimates from dichotomous choice contingent valuation studies', *Land Economics,* vol. 67, pp. 64–73.

Ready, R.C. and Hu, D. 1995, 'Statistical approaches to the fat tail problem for dichotomous choice contingent valuation', *Land Economics,* vol. 71, pp. 491–9.

Sellar, C., Chavas, J.-P. and Stoll, J.R. 1986, 'Specification of the logit model: the case of valuation of nonmarket goods', *Journal of Environmental Economics and Management,* vol. 13, pp. 382–90.

Sheppard, R., Kerr, G.N., Cullen, R. and Ferguson, T. 1993, *Contingent Valuation of Improved Water Quality in the Lower Waimakariri River,* Research Report No. 221, Agribusiness and Economics Research Unit, Lincoln University, New Zealand.