Paper to be presented at the CES/Health Economics Study Group, Paris, 14th–16th January 2004.

Work in progress. Please do not quote or cite without permission

Testing for Inconsistencies in Double Bounded Dichotomous Choice Contingent Valuation Studies

Verity Watson (v.watson@abdn.ac.uk) and Mandy Ryan (m.ryan@abdn.ac.uk)
Health Economics Research Unit
Institute of Applied Health Sciences
University of Aberdeen

A number of issues have been raised in the contingent valuation (CV) literature regarding the 'validity' of CV welfare estimates. This study addressed a number of these concerns within a study looking at the value of an air ambulance service in England. The double bounded dichotomous choice (DBDC) CV approach was used, and data was collected from 800 individuals across England, using a computer assisted telephone interview. Consideration was first given to whether preferences were a function of the sequence of questions. There was evidence of this. We then consider a number of possible causes for such inconsistencies, including *framing, cost-based responses, strategic behaviour, yea-saying*, *and anchoring.* These were investigated by considering the effect of path dependency on the responses to the DBDC bids, where responses to 'ascending' (positive responses) and 'descending' (negative responses) sequences for a number of starting points are compared. Suggestive evidence of cost-based responses, strategic behaviour and yea-saying is provided. Areas for future research are identified.

Acknowledgements: Financial support from the Department of Health, University of Aberdeen and Health Foundation is acknowledged. The Chief Scientist Office of the Scottish Executive Health Department (SEHD) funds HERU. The views expressed here are those of the authors and not necessarily those of SEHD.

1. Introduction

The last 10 years has seen an increase of research into methods of eliciting a monetary valuation for the benefits from health care. Within health economics, this has been partly due to the need for a method of benefit assessment that goes beyond health outcomes. Within environmental economics, and some sections of the health economic community, this research has also been driven by a desire for a monetary measure of benefit that can be directly compared with costs within the framework of an economic evaluation. Given the lack of a market for many health and environmental goods, the contingent valuation method (CVM) has been used to directly elicit willingness to pay (WTP) from respondents for a specified change in resource allocation for non-market goods (Carson 2003; Ryan et al, 2001). Despite the increased use of the CVM, there is still widespread debate over the validity of the technique and, related to this, the choice of elicitation format.

There is some degree of consensus across both disciplines that the open-ended approach is not appropriate (Donaldson et al 1997a; Arrow et al 1993; Ryan et al, 2004). Within health care, the payment card approach has proved popular. This technique presents respondents with a range of bids and respondents circle the amount that represents the most they would be willing to pay. Donaldson et al (1997a) argued that this method also mimics real life by allowing individuals to "shop around" for the value that is the most they would. Comparisons of the PC and OE questions show that respondents are more likely to answer PC questions, that more consistent mean and median values are obtained by the use of a PC and that a stronger association between WTP and ability to pay is obtained (Donaldson *et al*., 1997a) leading to this technique being used in healthcare (Donaldson *et al.,* 1995; Ryan, 1996; Donaldson and Shackley, 1997; Donaldson *et al.,* 1997a; Donaldson *et al.,* 1997b; Gibb *et al.,* 1998; Olsen and Donaldson, 1998)

Since the recommendations of the NOAA panel, the dichotomous choice (DC) method has proved popular (Johannesson *et al*., 1991; Johannesson *et al.,* 1993; Ryan, 1997; Ryan, 1998; Frew et al, 2003; Ryan et al, 2004). This approach asks individuals whether they would pay a specified amount for a given commodity, with possible responses usually being "yes" or "no". The bid amount is varied across respondents and the only information obtained from each individual is whether his/her maximum WTP is above or below the bid offered. The main

arguments for its use are simplicity for respondents and incentive compatibility resulting in reduced incentives for strategic behaviour (Carson et al, 1999; Hoehn and Randall, 1987). However, the DC method does have the drawback that it provides limited information for an individual. Thus, for a researcher to obtain a given level of statistical precision DC studies require larger sample sizes than the open-ended and payment card approaches.

To increase the information obtained from the initial dichotomous choice question (DC1), researchers can add a subsequent dichotomous choice question (DC2) resulting in a double bounded dichotomous choice (DBDC) question format. An individual's response to the DC1 question determines the bid level offered in DC2. Where, if respondents answer 'yes' to the DC1 question a higher bid level is offered in DC2. Conversely, if respondents answer 'no' to DC1 they are offered a lower bid level in DC2. Hanemann et al (1991) found that DBDC data reduced the variance of WTP relative to single bounded dichotomous choice.

However, welfare estimates generated from the DBDC format have been found to be lower than those estimated from the single DC question. Hanemann, et al (1991) found that the DBDC format reduced the point estimates of WTP relative to the single bounded format. Bateman et al (2001) conducted a similar analysis for multiple bounded dichotomous choice elicitation format. They found welfare estimates decreased over bounds, the largest decrease being observed between the first and second bound (the initial and follow up bid). Bateman et al suggest that their result is consistent with their indignation/guilt hypothesis, resulting from the iterative process involved in DBDC. Indignation and guilt are two sides of the same effect applying to the bid-increasing path and the bid-decreasing path respectively. Indignation will occur when respondents perceive that in answering DC1 they have struck a deal with the interviewer. Thus when asked the higher amount in DC2 respondents feel that the interviewer has gone back on the deal. This increases the likelihood of a negative follow up response. The guilt hypothesis occurs when respondents who state 'no' to DC1 are asked a lower bid amount in DC2. This lower amount makes respondents feel a sense of social responsibility or guilt. Thus, the guilt hypothesis will increase the probability of a positive response to follow up questions. Alberini et al (1997) and Cameron and Quiggan (1994) found that the relationship between WTP and socio-economic determinants of WTP differed between the analysis of DC1 and DC2, where responses to DC2 were dependant on DC1.

Given the hypothetical nature of CV markets a key issue concerns the validity of the WTP estimates (whatever method of elicitation is chosen). A number of concerns have been expressed in the literature regarding the application of the CVM to value non-market commodities. These include: framing; cost-based responses (where respondents try to estimate the cost of the commodity rather than give their value; strategic behaviour (where respondents deliberately misrepresent their true preferences to influence future payment for or provision of the good); yea-saying (a tendency for respondents to agree with the interviewer irrespective of the true valuation they place on the good or service) and anchoring (when the initial bid level is seen to conveys information to the respondent) (Ryan et al, 2001).

The aims of this paper are two-fold. Firstly, the paper considers the application of the DBDC format for estimating WTP in health care. More specifically, consideration is given to whether the findings in the environmental literature regarding WTP estimates from the DBDC approach being lower than that from the single bounded question are observed when this technique is applied to health care. If this inconsistency is found potential reasons are explored. Section 2 presents an overview of data used for this study, collected from a telephone survey of the general population of England to consider the value of air ambulance provision. The policy context is community preferences for air ambulance services. Whilst the results presented here provide some insight as to the value that the community places on air ambulances, the primary objective is to investigate anomalies in WTP data. Section 3 directly tests whether the DBDC approach generates inconsistent responses in terms of WTP estimates varying according to the initial and subsequent response. Section 4 further investigates the DBDC responses. We consider a number of possible causes for such inconsistencies, including *framing*; *cost-based responses; strategic behaviour; yea-saying* and *anchoring.* These were investigated by considering the effect of path dependency on the responses where responses to 'ascending' (following a positive response to initial bid) and 'descending' (following a negative response to initial bid) sequences within the questionnaire for a number of starting points are compared. This is very much work in progress and all comments will be gratefully received.

2. Application to valuing Air Ambulances

Since air ambulances were first introduced in England the number has risen to 14 providing full national coverage, this increase was encouraged by the establishment of the National Association of Air Ambulance Services (NAAAS) in 1997 and commercial sponsorship amounting to £14 million over three years by the AA Foundation in 1999. This commercial sponsorship funded the aircraft related costs and contributed to financing additional aircraft, whilst the NHS financed the provision of clinical staff. Air ambulances have a recognised role in providing emergency care however; recent research has questioned whether they are an optimal use of NHS resources. This has become a greater consideration since AA Foundation sponsorship ceased in January 2002, with no alternative sponsor being identified.

This study, funded by the Department of Health, sought public perceptions of and WTP for the provision of a national air ambulance service, to inform decision-making regarding the future funding of an air ambulance service. Between August and September 2002, a representative sample of 1400 members of the public was interviewed by using computer assisted telephone interviews (CATI). For results on the general survey see Johnston and Ryan (2002).

A DBDC elicitation format was used to elicit maximum WTP. Respondents were randomly presented with one of five possible 'base bids': £25, £50, £100, £200, £300. If a 'Yes' ('No') response was given to the first WTP question, respondents were offered a higher (lower) follow up WTP question. The follow up questions are as stated for each bid level; £25 (with lower follow up of £10 and higher follow up of £50); £50 (with lower follow up of £25 and higher follow up of £100); £100 (with lower follow up of £50 and higher follow up of £200); £200 (with lower follow up of £100 and higher follow up of £300); and £300 (with lower follow up of £200and higher follow up of £400). The bid levels were taken from a willingness to pay survey of air ambulances conducted in the Grampian Region of Scotland (Ryan et al, 2004). Responses to these bid amounts were monitored throughout the data collection period. Table 1 below presents the data collected in the survey. Intuitively expectations are that as the bid levels increase the probability of acceptance should fall. This pattern is observed for the data.

Table 1: Responses to DBDC bid levels

'Base bid'	Yes	No	Upper bid	Yes (Upper)	No (Upper)	Lower bid	Yes (lower)	No (Lower)
£25	68.3	14	£50	75	24	£10	35	65
£50	65.3	18.19	£100	65	29	£25	34	66
£100	52	45	£200	55	45	£50	21	78
£200	37	62	£300	44	55	£100	23.2	76
£300	27	76	£400	21	78	£200	20	79

3. Testing for consistency in DBDC responses

3.1 Methods

Prior to respondents being asked the DBDC questions they were initially asked a if they would be willing to pay anything for an air ambulance service. If they answered 'yes' then they were offered a randomly assigned initial bid (DC1). If respondents stated 'no' they were asked why. Given that the main objective of this a paper is to consider anomalies in DBDC, respondents who stated 'no' to the screening question are not considered in the analysis presented in this paper.[1] 802 respondents stated that they would be willing to pay something for an air ambulance service.

Moving to a DBDC data format from single bounded dichotomous choice question increases the complexity of the data analysis (Haab and McConnell, 1993). Initially a general model, the bivariate probit is used to analysis the data (Model I).The results of two separate probit models, one for DC1 (Model II) and one for DC2 (Model III), are then presented. Here the data is analysed as though it was generated by two different single bounded dichotomous choice experiments. This allows the consideration of the well documented results from environmental economics that DBDC results in lower welfare estimates than single bounded. In addition, an interval data model is estimated as proposed by Hanemann et al (1991) where this is a special case of the bivariate model (Model IV). In general let t_1 be the first bid price and t_2 be the second. Then:

A Yes | Yes response ↵ WTP □ t_2
A Yes | No response ↵ t_1 □WTP < t_2
A No | Yes response ↵ t_1 > WTP □ t_2
A No | No response ↵ WTP < t_2

[1] In the valuation exercise those who answered no to the screening question were asked why – the aim here was to distinguish 'protestors (who need to be excluded from the analysis) from genuine zero values (who need to be included in WTP estimations, Johnston and Ryan, 2002).

Model I – Cameron and Quiggan (1994) argued that the responses to the two DC questions may not be independent, but rather path dependent i.e. the response to DC2 may be related to the response to DC1. Thus, they proposed the use of a more general model in the analysis of DBDC data, using the bivariate model approach where:

$$y_{i1}^{*} = X_{i1}\beta_1 + \varepsilon_{i1}$$

$$y_{i2}^{*} = X_{i2}\beta_2 + \varepsilon_{i2}$$

in the data y_{i1} and y_{i2} are only observable through individuals responses to DC1 and DC2 respectively such that:

$$y_{i1} = \begin{cases} 1 \text{ if } y_{i1}^{*} > t_1 \\ 0 \text{ otherwise} \end{cases} \text{ and } y_{i2} = \begin{cases} 1 \text{ if } y_{i2}^{*} > t_2 \\ 0 \text{ otherwise} \end{cases}$$

The error terms ε_{i1} and ε_{i2} are unobservable components of WTP distributed $N(0,\sigma_1^2)$ and $N(0,\sigma_2^2)$ respectively. These unobservable components may take two forms, those that are individual specific and the same across equations ε_i and those that are individuals specific and differ between the valuation questions ε_{i1} and ε_{i2}. In the case of DBDC data there are two discrete responses I_{i1} and I_{i2}. As the bid amount offered in DC2 is dependant on the response to DC1 the model must be developed in the context of the joint distribution of (y_{i1} and y_{i2}). In this case Cameron and Quiggan proposed a bivariate normal distribution $BVN(x_1'\beta_1,\ x_2'\beta_2,\ \sigma_1^2,\ \sigma_2^2,\ \rho)$ for both valuations, where ρ is the correlation parameter for the equations. If $\rho<1$ it implies that there is a question specific effect in the error term, where respondents do not view DC1 and DC2 as being identical.

Models II and III - responses to the DC1 (Model II) and DC2 (Model III) are analysed separately, as though they were elicited from two separate independent dichotomous choice experiments. In this case there is no correlation between the responses thus assuming ρ =0. If there is no correlation between DC1 and DC2 then joint estimation would provide no statistical gain.

Model IV - Following the method proposed by Hanemann et al (1991) an interval data model is used. Here it is assumed that responses to both valuation questions, DC1 and DC2, are motivated by the same latent WTP value. Thus, the covariates and the errors for when analysing both dichotomous choice questions are identical. It should be noted that the model proposed by Hanemann et al (1991) is a special case of the bivariate model where:

$$X_{i1} = X_{i2}$$
$$\beta_1 = \beta_2$$
$$\sigma_1 = \sigma_2$$

and ρ =1 implying that there is perfect correlation between responses to DC1 and DC2. Where, individuals' valuations of the good are identical in DC1 and DC2 with any observed difference due to randomness in the underlying WTP.

For simplicity, for all models estimated the only independent variables included are the constant and the bid vector.

3.2 Results

Table 2 below reports the results. In Model I, the bivariate probit, the first four rows report the coefficients and standard errors for responses to DC1 and the second four rows are the coefficients and standard errors for responses to DC2. In model I ρ is positive and significantly different from zero, thus implying a positive correlation between responses to DC1 and DC2. As $\rho < 1$ it implies that although ε_{1i} and ε_{2i} are correlated they are not identical. In considering the results of Model II, in comparison with Model IV, it can be seen that findings first reported by Hanemann et al (1991) that DBDC reduces the point estimates

of WTP relative to single bounded dichotomous choice hold for this data set. The results of Models II and III suggest that as respondents move from the initial bid (DC1) to the follow up bid (DC2) they are more likely to reject the WTP amount, the consequence of this is that WTP estimates fall dramatically from the initial bid analysis to the follow up bid analysis. These results are consistent with those findings from Bateman et al (2001). The results of these models indicate that the inconsistencies found to be present in the environmental literature are found in the case of this data set.

Table 2: Welfare estimates from different models

		Model I Bivariate Probit Model	Model II First Bid Responses Only	Model III Second Bid responses Only	Model IV Interval Data Model
Constant		0.7426	1.1816	0.2013	1.6528
		(0.0785)**	(0.1294) **	(0.1276) **	(0.1097)**
Bid		-0.0052	-0.0084	-0.0061	-0.0158
		(0.0005)**	(0.0007) **	(0.0009) **	(0.0008) **
Upper Bound			.	.	.
Constant		-0.0507	.	.	.
		(0.0917)	.	.	.
Follow up		-0.0024	.	.	.
		(0.006)**	.	.	.
Observations		802	802	802	802
Log Likelihood		984.7121	491.1251	497.75472	1224.4172
Rho		-0.2389			
		(0.0771) **			
Mean WTP	Lower Bound	£142.81	£141.05	.	.
	Upper Bound	-£21.13	.	£33	.
	Overall		.	.	£104.34

** denotes significance at the 1% level
Standard errors are reported in parenthesis.

4 Testing for inconsistencies in WTP responses

4.1 A priori hypotheses

Section 3 provides evidence that supplementing the initial question of a DBDC question with a subsequent dichotomous question generates inconsistent responses. That is, there seems to be evidence that the distribution of underlying preferences implied by the initial question are

not the same as that implied when analysing both questions. In this section we consider a number of possible causes for this finding. Following Deshazo (2002), responses are divided into ascending (resulting from a positive initial response) and descending (resulting from a negative initial response) paths. Expected patterns of responses, corresponding to different potential causes of inconsistencies, are outlined in Table 3.

Table 3: Testing for inconsistencies in DBDC responses: a priori hypotheses

Interval	Bid Increasing Path	Expected Outcome	Bid Decreasing Path
Reference dependent preferences/framing			
Upper	P(Y | Y)	>	P(Y)
Middle	P(Y | N)	>	P(N | Y)
Lower	P(N)	=	P(N | N)
Cost-based responses			
Upper	P(Y | Y)	<	P(Y)
Middle	P(Y | N)	>	P(N | Y)
Lower	P(N)	<	P(N | N)
Strategic behaviour			
Upper	P(Y | Y)	<	P(Y)
Middle	P(Y | N)	>	P(N | Y)
Lower	P(N)	>	P(N | N)
Yea-saying			
Upper	P(Y | Y)	>	P(Y)
Middle	P(Y | N)	<	P(N | Y)
Lower	P(N)	>	P(N | N)
Anchoring			
Upper	P(Y | Y)	<	P(Y)
Middle	P(Y | N)	?	P(N | Y)
Lower	P(N)	>	P(N | N)

Reference Dependant Preferences and Prospect Theory / Framing

Prospect theory, as developed by Kahneman and Tversky (1979), proposes that in a choice setting individuals will form a reference point with probability p. In subsequent choice settings the individual will frame outcomes as a gain or a loss against this reference point. This implies that if equivalent respondents have different probabilities of forming reference points then it would appear that they respond differently to equivalent WTP questions. DeShazo (2002) applies this theory to the context of DBDC format. If respondents answer ‘yes’ to the initial bid

they form a reference point at this level. However, if respondents answer ‘no’ to the initial bid level no reference point is formed. Resultantly, when the respondents who answer ‘yes’ to the initial bid compare their reference point to the higher follow up bid, this outcome is framed negatively.

The framing explanation, based on prospect theory, suggests that respondents form a reference point at the DC1 bid level. Thus, a bid-increasing path negatively frames DC2, as it is higher than DC1. Under this explanation, the proportion of respondents answering ‘yes’ to the negatively framed DC2 question, P(Y|Y), will be less that than the probability of respondents answering 'yes' to the unframed equivalent P(Y) in a bid decreasing path. The respondents who answered no to the negatively framed second bid will end up in the P(Y|N) interval where this will be larger than the equivalent interval in the bid-decreasing path P(N|Y). Finally DeShazo (2002) predicts that the proportion of respondents in the lower intervals P(N) and P(N | N) in the bid-increasing and bid-decreasing paths respectively will be equal.

Cost Based Responses

A number of CV experiments have found evidence of cost-based WTP responses. Schkade and Payne (1994), in using verbal protocol analysis to investigate how people respond to WTP questions, found that individuals justify their WTP responses by referring to the cost of the commodity being valued. Donaldson et al (1997a) found evidence of cost based responses in a study looking at different methods of testing for cystic fibrosis. They argued that this problem may arise when open-ended questions are used to elicit WTP values. However, a study by Ryan and San Miguel (2000) found evidence of cost-based responses when using the payment card format.

Within the context of the DBDC format, and developing the cost-based responses argument, Carson et al (1994) argued that a follow-up question might be seen as an attempt by the government to obtain additional funds, beyond the actual cost of the good. Thus, individuals will be more likely to say ‘no’ to a follow-up question. Alternatively, in the case of a lower follow-up, it may be perceived that a lower quality of good will be provided and again there is an increased probability that the respondent will say no.

. Applying the cost-based response argument, Carson et al (1999) proposed a number of explanations for the observed relationship between the single bounded dichotomous choice question and the DBDC approach. Using the hypothesis that additional bids invoke uncertainty about the cost, they put a number of arguments forward. Firstly, when faced with an initial scenario and DC1 respondents expect that the information presented is known with certainty, thus when DC2 is presented uncertainty over the cost of the good is created. The cost uncertainty explanation suggests that individuals respond to the uncertainty created by the follow-up by treating the responses to DC1 and DC2 as independent. Under the assumption that respondents are risk adverse there is an increased probability that respondents will answer 'no' to DC2. A further explanation is that responses to DC2 are not based on the bid offered in DC2 but on the weighted average of DC1 and DC2. Here respondents faced with a second bid level assume that the true cost of the good lies somewhere between DC1 and DC2. The consequence of this is an extreme response to DC2. For instance, if a respondents true WTP y* was such that DC1<y* < DC2 or DC1> y* >DC2 then responses of YES|NO and NO|YES should be observed. However if respondents perceive the true cost as being the weighted average of DC1 and DC2 it is more likely that responses will be YES |YES and NO | NO. Carson et al (1999) suggest that responses to the follow-up question are a form of bargaining where in the case where DC2 > DC1 the optimal response would be no. The observed impact on the data is the same as in the cost uncertainty hypothesis, only with a different psychological explanation. The final proposition is that the bid level conveys information to the respondents about the quantity or quality of the good provided. Thus, if as DC1 and DC2 differ respondents perceive they are being asked to differing goods. Consequently, when DC1 < DC2 the good offered at the second bid level is assumed to be of a higher quality or more is provided and vice versa. Here respondents in the upper interval will be lower in the bid-increasing path that the bid-decreasing path thus P(Y | Y) < P(Y). Due to a combination of the government wastage and the reduction in quality arguments, discussed above, the proportion in the middle interval will be higher in bid-increasing path P(Y | N) than a bid-decreasing path P(N | Y). Based on the reduction in quality argument respondents in the lower interval of a bid decreasing path P(N |N) will be greater that the bid-increasing path P(N).

Strategic Behaviour

Respondents are more likely to answer 'no' to any follow up question. This is a result of adding a follow up question to the dichotomous choice format where the follow up is believed to compromise the incentive compatibility of the elicitation method. The NOAA panel stated that if a elicitation format is not incentive compatible then respondents have an incentive to engage in strategic behaviour, such as stating ‘no’ in order to minimise the amount that they may have to pay.

Yea Saying

There is a general concern, raised in the psychological and sociological literatures, about the tendency of people towards “yea-saying” (i.e. answering “Yes” to a question) when this is thought to be the socially desirable response (Couch and Kenitson, 1960; Bachman and O’Malley, 1984). Applying this to CV experiments, it can be hypothesised that surveys using the DC approach would produce values which are biased in an upwards direction relative to some other method, such as the PC (or OE), for which it may be assumed that such yea-saying is not an issue. This hypothesis is supported by studies comparing the PC and OE techniques with the DC method (Bishop *et al.,* 1983; Sellar *et al.,* 1985; Johnson *et al.,* 1990; Boyle *et al.,* 1993; Kealy and Turner, 1993; McFadden, 1994; Kristrom, 1997; Ready *et al.,* 1996; Kramer and Mercer, 1997; Frew et al, 2003; Ryan et al, 2004)). With the exception of Kramer and Mercer, all of these studies have found that the DC method consistently gives higher welfare estimates than the OE or PC techniques. A number of authors involved in these studies have gone on to express concern about “yea-saying” within the context of DC questions (McFadden and Leonard, 1993; Boyle *et al.,* 1993; Kannimen, 1995; Holmes and Kramer, 1995; Ready *et al.,* 1996) whilst others have attempted to test for “yea-saying”. Kanninen (1995) developed a test to identify the number of “yea-saying” respondents in a DC survey and concluded that 20% of respondents could be identified as yea-sayers. Ready *et al.,* (1996) attempted to test for “yea-saying” and concluded that their findings were consistent with the hypothesis of “yea-saying”.

Applying the yea-saying hypothesis to DBDC format, it may be argued that: (i) yea-sayers tend to say “yes” to any bid offered; (ii) “yea-saying” should not affect the descending sequence because it starts with a “no” answer; and following on from (ii), (iii) there is no bias in the descending sequence. Thus, if “yea-saying” occurs, an upward bias will affect the ascending sequence, leading to an overall upward bias in WTP estimate. Thus, the descending sequence, assumed to be free from bias, can be used as a reference, and compared with the ascending sequence, which might be influenced by “yea-saying”.

Following this, the model predicts that in the upper WTP interval the proportion of respondents answering “yes” to DC1 and also “yes” to DC2 in the ascending sequence will be more than the proportion of those answering “yes” to DC1 in the descending sequence. In the middle interval, the proportion of respondents answering “yes” to DC1but then “no” to DC2 in the ascending sequence will be smaller than the proportion of those answering “no” to DC1 but then “yes” to DC2 in the descending sequence. In the lower WTP interval, the proportion of respondents answering “no” to DC1 in the ascending sequence will be greater than the proportion of those answering “no” to DC1 and then “no” to DC2 in the descending sequence.

Anchoring

In this case, the starting point serves as an anchor where respondents interpret this as the true value of the good. Thus $P(Y|Y) < P(Y)$ and $P(N) > P(N|N)$ but the effect of this upon the middle interval of the DBDC sequence is ambiguous.

4.2 Results

The results are shown in Table 4. It can be seen from the response patterns for Bid Groupings 1-3 the response patterns are consistent with explanations of cost-based responses or strategic behaviour. Additionally the patterns observed for bid groupings 1-3 would be consistent with the anchoring hypothesis - with this hypothesis the predictions regarding the effect of the middle interval is ambiguous. Considering bid grouping 2 the proportion of respondents in the lower interval for the bid increasing and bid decreasing paths are close - this may point to framing as an appropriate explanation.

Table 4: Results from tests for consistency of responses

	Bid Increasing Path	Proportion (%)	Bid Decreasing Path	Proportion (%)	Observed Pattern	Possible explanations
			Grouping 1: £25 and £50			
Upper Interval	$P(Y_{25} \mid Y_{50})$	14.95	$P(Y_{50})$	37.05	<	Cost based responses / Strategic behaviour
Middle Interval	$P(Y_{25} \mid N_{50})$	27.76	$P(N_{50} \mid Y_{25})$	11.15	>	
Lower Interval	$P(N_{25})$	13.16	$P(N_{50} \mid N_{25})$	5.76	>	
			Grouping 2: £50 and £100			
Upper Interval	$P(Y_{50} \mid Y_{100})$	12.59	$P(Y_{100})$	32.62	<	Cost based responses / Strategic behaviour
Middle Interval	$P(Y_{50} \mid N_{100})$	24.46	$P(N_{100} \mid Y_{50})$	15.6	>	
Lower Interval	$P(N_{50})$	16.91	$P(N_{100} \mid N_{50})$	12.77	>	
			Grouping 3: £100 and £200			
Upper Interval	$P(Y_{100} \mid Y_{200})$	7.09	$P(Y_{200})$	20.14	<	Cost based responses / Strategic behaviour
Middle Interval	$P(Y_{100} \mid N_{200})$	25.53	$P(N_{200} \mid Y_{100})$	15.11	>	
Lower Interval	$P(N_{100})$	28.36	$P(N_{200} \mid N_{100})$	19.06	>	
			Grouping 4: £200 and £300			
Upper Interval	$P(Y_{200} \mid Y_{300})$	4.68	$P(Y_{300})$	13.87	>	Yea - Saying
Middle Interval	$P(Y_{200} \mid N_{300})$	15.47	$P(N_{300} \mid Y_{200})$	9.96	<	
Lower Interval	$P(N_{200})$	34.17	$P(N_{300} \mid N_{200})$	37.37	<	

It can be seen from the table that the patterns observed from bid group 4 differ from the patterns in bid groups 1 –3. This bid pattern indicates that yea-saying may be an appropriate explanation for this pattern.

5. Tentative conclusions and areas for future research

The results of the empirical analysis in section 3 of this paper provide evidence that the anomalies reported in DBDC data from environmental economics are present in data collected in a health care context. That is, the welfare estimates derived from analysing the DBDC data are lower than those found if the data is considered as single bounded dichotomous choice. Additionally, the results of the bivariate probit model indicate that there is low correlation between DC1 and DC2. This questions that assumption that respondents reveal well-defined preferences when asked a WTP question. Based on these findings the paper moved on to consider possible explanations, following DeShazo (2002). As opposed to DeShazo (2002), who found evidence of framing in DBDC data, the results in this paper are consistent with cost based responses or strategic behaviour. At the highest initial bid level the pattern of responses suggests yea saying may exist.

One limitation of this analysis is that for certain response patterns several of the proposed explanations for anomalies apply. At present it is not possible to differentiate cost based responses from strategic bias. Supplementing the analysis presented in this paper with consideration of qualitative follow-up responses may form an explanation that is more precise and provide more insight.

When the NOAA panel recommended the use of dichotomous choice as an appropriate elicitation format it was based on the incentive compatibility of the format. In the move from a single to a double-bounded dichotomous choice structure, it is suggested that incentive compatibility is lost. One consequence of non-incentive compatibility is the presence of strategic bias, which was found in this study. Future work should incorporate a test of incentive compatibility, similar to that proposed by Whitehead (2002). If results indicate that anomalous results occur when the elicitation format is not incentive compatible, this has implication for not only the use of DBDC but also all elicitation formats (including the payment

card, which has been favoured in Health Economics) other than a single bounded dichotomous choice question.

In the case of respondents offered the highest initial bid level yea-saying was suggested by the response patterns presented in table 4. Further analysis may consider applying the methods adopted by Kanninen (1995) and Whitehead (2002) to test for yea-saying. Liljas and Blumenschein (2000) suggest that yea saying may occur when respondents are uncertain about their 'true' valuation or interpret the bid level they are presented with as being the cost of the good as it the case in cost based responses. Liljas and Blumenschein (2000) suggest that this may be reduced if responses are 'calibrated' by respondents' self-reported certainty. The air ambulance data reported here also collected certainty data, thus future work should consider if the observed anomalies are reduced if only respondents who stated they were 'certain' or 'very certain' are included in the analysis.

Considering the guilt hypothesis explanation for anomalies in the DBDC data, further work may explore if the choice of payment vehicle affects the strength of this effect on data sets. It maybe hypothesised that respondents feel a heightened sense of guilt or social responsibility if the payment vehicle chosen was donations to charity as opposed to increased taxation. Again, the data reported here asked WTP in terms of either increased taxation or a charitable donation so future work wil compare results.

Ultimately the motivation for valuing goods using stated preference methods is to obtain a value for a good where currently no market exists. This paper presents test of the theoretical validity of the DBDC elicitation format. Tthe validity of the format is questioned as it results in lower estimates of mean WTP than the single bounded DC format. Many preference elicitation formats maybe compared with one another, thus allowing hypotheses regarding convergent validity of methods to be made. However, the true test of a stated preference format is in a comparison with revealed preference valuation data for the same good. In this case the good discussed is the provision of an air ambulance service. In some counties of England local air ambulances are provided through charitable funds financed by the community. Further work should consider whether the results of this study are significantly different to the values implied by charitable donations to air ambulance funds.

References

Alberini, A., (1995), Optimal designs for discrete choice contingent valuation surveys: Single-bound, double-bound and bivariate models, *Journal of Environmental Economics and Management,* **28**, pp287-306.

Alberini, A., (1995), Efficiency vs Bias of willingness to pay estimates: bivariate and interval data models, *Journal of Environmental Economics and Management*, **29**, pp169-180.

Alberini, A. Kanninen B. and Carson R. (1997) Modelling response incentive effects in dichotomous choice contingent valuation data. Land Economics **73**, 309-324.

Altaf, M. A., (1994), Household demand for improved water and sanitation in a large secondary city, *Habitat International,* **18(1)**, pp45-55.

Arrow, K., Solow, R., Portney, P. R., Leamer, E. E., Radner, R., And Schuman, H., (1993), Report of the NOAA panel on contingent valuation, *Federal Register*, **58**, pp4601-4614.

Bachman, J.G., O'Malley, P.M., 1984. Yea-saying, nay-saying, and going to extremes: black-white differences in response styles. Public Opinion Quarterly, 48, 491-509.

Bateman, I. J., Landford, I. H., Jones, A. P., and Kerr, G. N., (2001), Bound and path effects in double and triple bounded dichotomous choice contingent valuation, *Resource and Energy Economics,* **23**, pp191-213.

Bishop, R., Heberlein, T., Kealy, M., 1983. Contingent valuation of environmental assets: comparisons with a simulated market. Natural Resources Journal*,* 23, 619-633.

Blamey, R. K., Bennett, J. W., and Morrison, M. D., (1999), Yea-saying in contingent valuation surveys, *Land Economics*, **75(1),** pp126-141.

Boyle, K. J., Experimental designs affect preference measures, ????

Boyle, K., Johnson, F., McCollum, D., Desvousges, W., Dunford, R., Hudson, S., 1993. Valuing public goods: discrete versus continuous contingent-valuation responses. Department of Resource Economics, Working Paper, University of Maine.

Burton, A. C., Carson, K. S., Chilton, S. M., Hutchinson, W. G., (2003), An experimental investigation of explanations for inconsistencies in responses to second offers on double referenda, *Journal of Environmental Economics and Management,* Forthcoming.

Couch, A., Kenitson, K., 1960 Yeasayers and naysayers: agreeing response set as a personality variable. Journal of Abnormal and Social Psychology, 60, 151-174.

Cameron, T., and Quiggan, J, (1994), Estimation using contingent valuation data from dichotomous choice with follow up questionnaire, *J Environ. Econom. Management,* **27** pp218-234.

Carson, R. T., (2003), *A Bibliography of Contingent Valuation Studies,* Edward Elgar, Cheltenham.

Carson, R. T., Mitchell, R. C., Hanemann, W. M., Kopp, R. J., Presser, S., and Ruud, P. A., (1994), Contingent valuation and loss passive use: damages from the Exxon Valdez, *Discussion Paper, 94-18,* Resources for the Future, Washington D.C.

Carson R. Groves T. Machina M. 1999. Incentive and informational properties of preference questions – plenary address. In: Proceedings of the Ninth Annual Conference of the European Association of Environmental and Resource Economics (EAERE), Oslo, Norway, June 1999.

DeShazoJ. R., (2002), Designing Transactions without framing effects in iterative question formats, *Journal of Environmental Economics and Management*, **43**, pp360-385.

Donaldson, C., Shackley, P., Abdalla, M., Miedzybrodzka, Z., (1995), Willingness to pay for antenatal carrier screening for cystic fibrosis, *Health Economics*, **4**, pp439-452.

Donaldson, C., Shackley, P., (1997), Does process utility exist? A case study of willingness to pay for laparoscopic cholecystectomy, *Social Science and Medicine*, **44**, pp699-707.

Donaldson, C., Thomas, R., Torgenson, D., (1997a), Validity of open-ended and payment scale approaches to eliciting willingness to pay, *Applied Economics*, **29**, pp79-84.

Donaldson, C., Shackley, P., Abdalla, M., (1997b). Using willingness to pay to value close substitutes: carrier screening for cystic fibrosis revisited, *Health Economics*, **6**, pp145-159.

Frew, E.J, Whynes, D.K., Wolstenholme, J.L., 2003. Eliciting willingness to pay: comparing closed-ended with open-ended and payment scale formats. Medical Decision Making, 23, 150-159.

Gibb, S., Donaldson, C., Henshaw, R., (1998), Assessing strength of preference for abortion method using willingness to pay, *Journal of Advanced Nursing,* **27**, pp30-36.

Haab, T. C., and McConnell, K. E., (2003*), Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation*, New Horizons in Environmental Economics, Edward Elgar.

Hanemann, W. M., Loomis, J., and Kanninen, B., (1991), Statistical Efficiency of Double Bounded Dichotomous Choice contingent valuation, *Amer. J. Agri. Econom.* **73**, pp1255-1263.

Hoehn J. and Randall A. 1987. A satisfactory benefit cost indicator from contingent valuation. Journal of Environmental Economics and Management, **14(3),** 112-131.

Holmes, T. P., and Kramer, R. A., (1995) An independent sample test of yea saying and starting point bias in dichotomous choice contingent valuation, *Journal of Environmental Economics and Management*, **29**, pp121-132.

Johannesson, M., Jonsson, B., Borgquist, L., (1991), Willingness to pay for antihyperintensive therapy – results of a Swedish pilot study, *Journal of Health Economics*, **10**, pp461-74.

Johannesson, M., Johansson, P.O., Kristrom, B., Gerdtham, U.G., (1993), Willingness to pay for antihypertensive therapy - further results, *Journal of Health Economics*, **12**, pp95-108.

Johnson, R.L., Bregenzer, N.S., Shelby, B., 1990. Contingent valuation question formats: dichotomous choice versus open-ended responses, in RL Johnson and GV Johnson (eds) Economic Valuation of Natural Resources: Issues, Theory and Applications, Boulder: Westview Press, pp. 193-204.

Johnston D. and Ryan M. (2002) Air Ambulances: public perceptions of value. Report prepared for Department of Health.

Kanninen, B., 1995. Bias in discrete response contingent valuation. Journal of Environmental Economics and Management, 28, 114-125.

Kealy, M., Turner, R., 1993. A test of equality of closed-ended and open-ended contingent valuations. American Journal of Agricultural Economics, 75, 321-331.

Kramer, R.A., Mercer, D.E., 1997. Valuing a global environmental good: U.S. residents' willingness to pay to protect tropical rain forests, Land Economics, 73, 196-210.

Kriström, B., 1997. Spike models in contingent valuation, American Journal of Agricultural Economics, 79, 1013-1023.

Liljas, B., and Blumenschein, K., (2000), On hypothetical bias and calibration in cost-benefit studies, **52**, pp 53-70.

McFadden, D., 1994. Contingent valuation and social choice. American Journal of Agricultural Economics, 76, 689-708.

McFadden, D., Leonard, G., 1993 Issues in the contingent valuation of environmental goods: Methodologies for data collection and analysis, in Contingent Valuation: A Critical Assessment, Cambridge Economics, Inc., Cambridge, MA.

McLeod D., And Bergland, O., (1999), Willingness to pay estimates using a double bounded dichotomous choice contingent valuation format: A test for validity and precision in a bayesian framework, *Land Economics*, **75(1)**, pp115-125.

Olsen, J.A., Donaldson, C., (1998), Helicopters, hearts and hips: using willingness to pay to set priorities for public sector health care programmes, *Social Science and Medicine,* **46**, pp1-12.

Ready, R.C., Buzby, J.C., Hu, D., 1996. Differences between continuous and discrete contingent value estimates, Land Economics, 72, 397-411.

Ryan, M., (1996), Using willingness to pay to assess the benefits of assisted reproductive techniques, *Health Economics*, **5**, pp543-558.

Ryan, M., (1997), Should government fund assisted reproductive techniques? A pilot study using willingness to pay, *Applied Economics,* **29**, pp841-849.

Ryan, M., (1998), Valuing psychological factors in the provision of assisted reproductive techniques using the economic instrument of willingness to pay, *Journal of Economic Psychology,* **19**, pp179-204.

Ryan, M., and San Miguel F., (2000), Testing for consistency in willingness to pay experiments, *Journal of Economic Psychology,* **21**, pp305-317.

Ryan, M., Scott, D. A., and Donaldson, C., (2004), Valuing health care using willingness to pay: A comparison of the payment card and dichotomous choice methods, *Journal of Health Economics*, Forthcoming.

Ryan M. Scott D A. Reeves C. Bate A. van Teijlingen E. Russell E. Napper M. Robb C. (2001) Eliciting public preferences for health care: a systematic review of techniques. *Health Technology Assessment* **5(5).**

Schkade, D., and Payne, J., 1994. How people respond to contingent valuation questions: A verbal protocol analysis of willingness to pay for an environmental regulation. Journal of Environmental Economics and Management **26**, 88-109.

Sellar, C., Chavas, J.P., Stoll, J.R., 1985. Validation of empirical measures of welfare change: a comparison of nonmarket techniques. Land Economics, 61, 156-175.

Whitehead, J. C., Clifford, W. B., and Hoban, T. J., (2000), WTP for research and extension programmes: Divergent validity of contingent valuation with single and multiple bound questions, Working paper, Department of Economics, East Carolina University.

Whitehead, J. C., (2002), Incentive compatibility and starting point bias in iterative valuation questions, *Land Economics*, **78(2)**, pp285-297.