

# Choice Rationality in Stated Preference Methods Applied to Farm Animal Welfare Improvements

Diane Burgess<sup>a\*</sup>, W. George Hutchinson<sup>b</sup>, Trevor McCallion<sup>c</sup>, Riccardo Scarpa<sup>d</sup>

a. Centre for Social and Economic Research on the Global Environment, University of East Anglia, Norwich, Norfolk, UK. NR4 7TJ. \* Corresponding Author, Tel: 01603 593741. E-mail: Diane.Burgess@uea.ac.uk

b. Department of Agricultural and Food Economics, Queens University Belfast, Newforge Lane, Belfast, UK. BT9 5PX

c. Biometrics Division, Department of Agriculture and Rural Development (Northern Ireland), Belfast, UK. BT9 5PX.

d. The Environment Department, University of York, York, UK. YO10 5DD

We examine the rationality of Stated Preferences for improving the welfare of farm animals by testing for preference reversal between choosing and pricing methods as found for risky choices by Tvesky et al (1990). To do this we use two independent methods for modelling stated preferences: Discrete Choice Contingent Valuation and Paired Comparisons.

In assessing whether the public is capable of rational decisions over highly complex and emotive public goods, the study determines the level of intransitive choices and examines for evidence of preference reversals in the estimated orderings provided by the methods. From a field survey of 400 respondents, we find a low level of intransitivity together with little evidence of preference reversals. It is concluded that the public possesses well-defined preferences and express consistent and transitive social choices between farm animal welfare schemes. Stated Preference approaches to Cost-Benefit analysis of animal welfare improvements appear relatively free from these types of preference anomaly.

## 1 Introduction

A prime objection to the use of Stated Preference (SP) Valuation Methods (eg Mitchell and Carson, 1989; Adamowicz *et al*, 1994) is the possible reversal of SP findings according to whether a choosing or pricing format is employed (e.g. Irwin *et al*, 1993; Brown, 1984). Within SP methods, an implicit assumption of rational choice is made invoking the axioms of utility theory, in particular those of transitivity, monotonicity and procedural invariance. The latter holds that preferences over alternatives are independent of the means of elicitation. Thus, individuals are characterised "*as having fully-formed and highly articulated preferences which they can quickly and accurately access and apply to any form of decision problem*" (Loomes, 1998, p478).

However, evidence of violations of the axioms of rational choice have been observed in people's behaviour. Tversky (1969) identified systematic, consistent and predictable intransitive preferences, while the preference reversal phenomenon was first observed by Lichtenstein and Slovic (1971) and Lindman (1971) where different method of elicitation: pricing and choosing produced different orderings of preferences for two binary lotteries:

- (1) P-bet - high probability, low payoff
- (2) \$-bet - low probability, high payoff

In these experiments, subjects were observed to reverse their preferences by choosing the "P-bet" yet valuing the "\$-bet" higher. These inconsistencies were suggestive of no optimising principles underlying the simplest of choices (Grether and Plott, 1979). Subsequent experiments were conducted to assess whether preference reversals occur in situations where economic theory is generally applied and whether it could be explained by economic theory (Grether and Plott, 1979, Pommerehne *et al*, 1982, Reilly, 1982 and Berg *et al*, 1985). These studies showed preference reversals to be a robust and persistent phenomenon.

In the most recent research, the Contingent Weighting model of Tversky *et al.* (1988) has been indicated as the cause of the observed choice-price discrepancy (Camerer, 1995). Here the importance (weight) attached to the attributes of a decision is deemed to be contingent on the method of elicitation. The observed choice-price discrepancy is induced by the compatibility principle which states “*the weight of any input component is enhanced by its compatibility with the output*” (Tversky *et al.* 1988, pg 376). In pricing decisions over the gambles, the pay-offs become more important than the probabilities as both the pay-offs and the values attributed to the gambles are expressed in monetary units. This causes the pay-offs to be overweighed in the pricing decisions resulting in the observed preference reversals between pricing and choosing elicitation methods.

The implicit implication of the occurrence of preference reversals is that instead of possessing fully formed preferences, people have imprecise preferences. From these, individuals will ‘construct’ their preferences using various rules of thumb, drawing on the salient characteristics and cues suggested by the elicitation method (Loomes, 1998, 1999). Thus, the preferences are likely to be affected by the framing of the context of the problems, as suggested by the prospect theory of Kahneman and Tversky (1979).

With the notion that people construct their preferences during the decision making process rather than possessing clear and well defined preferences, the implications for SP methods and more importantly the use of their output within policy-making are considerable. As SP methods are survey-based and therefore potentially susceptible to the framing of the survey questions, the question must be asked regarding whether information derived from the public regarding preferences and values used in policy decisions reflect as much (if not more) about the elicitation method as it does about the “true” preferences of the public? (Loomes, 1999). If by simply changing the method of elicitation causes the sign of the benefit-minus-cost figure to be reversed or substantially changed, this calls into doubt the results of cost-benefit analyses used for much policy analysis (Grether and Plott, 1979).

With particular regard to the method of contingent valuation (CV), preference reversals have been observed within a joint CVM and choice-based studies by Irwin *et al* (1993) and Brown (1984). Furthermore, the findings of Cox and Grether (1996) and Chu and Chu (1990) showed that persistent preference reversals could be eradicated within a market setting involving a number of arbitrations. As CV is generally employed to value non-market goods where arbitrage is not possible, preferences elicited within a CV have been shown to be highly susceptible to reversals. Thus, Irwin *et al* (1993, pg 17) states that “*a CV survey becomes an active process of value construction, rather than a neutral process of value discovery*”.

In recent years, a number of contingent valuation studies have been conducted to determine the benefits of improving farm animal welfare legislation (e.g. Bennett and Larson, 1996; Bennett and Blaney, 2003; Costello, 2000, (Glass 2001); Frykblom and Andersson, 2002; and Rolfe, 1999). Substantial estimates of WTP have been produced by these studies showing public support for improving the welfare of farm animals - providing support for the legislative changes that have been made at both the UK and European level<sup>1</sup> (Blandford and Fulponi, 1999). However, during the process of valuing the welfare improvement schemes, are the respondents acting as economically rational agents, or are they irrational and emotional in their stated behaviour? Are they able to express consistent preferences between welfare improvement programmes for different farm-animals? This is especially pertinent given the findings of preference reversals involving contingent valuation.

In this paper we examine whether the preference reversals found by economic psychologists and experimentalists also apply to technically complex and emotive choices between animal welfare improvement programmes. This is one of the first papers of its type to provide evidence on consistency between an SP pricing method, Contingent

---

<sup>1</sup> i.e. Council Directive 1991/629/EEC for the protection of calves (banning veal crates), Council Directive 91/630/EEC for the protection of confined pigs (banning sow stalls and tethers) and Council Directive 1999/74/EC for the protection of laying hens (banning convectional battery cages).

Valuation (CV) and an SP choosing method, Paired Comparisons (PC) when both are conducted in a large scale field survey over multiple goods.

The paper is organised as follows. In section 2 we outline a method for conducting both a choosing and pricing elicitation method over the same goods to enabling the potential preference reversals to be observed. Formal testing hypothesis are then stated. In Section 3, we outline sampling and analytical methodology introducing a novel Monte Carlo based approach to allow testing of the ordering obtained by each of the elicitation methods. Section 5 reports results while Section 6 discusses the implications of these findings and concludes.

## 2. The Research Design

We investigate whether intransitivity and preference reversals affect the SP findings in our CV study of WTP over four animal welfare improvement schemes<sup>2</sup>. To do so we compare our CVM field results with field results for the ordering of the same schemes based on the purely choice based method of Paired Comparisons (PC). This allows us to test whether the class of good in question (animal welfare improvements) and the type of discrete binary choice elicitation method favoured in current SP research (see Carson *et al* 1999) are subject to preference reversal. Such reversals have typically been found by economic psychologists when ordering risky choices using simple ranking and open ended valuation methods (e.g. Irwin *et al* 1993). These methods are more subject to strategic response than the discrete choice methods currently used in applied SP surveys, (See Carson *et al* 1999) Within the study, the respondents were first introduced to the four welfare improvement schemes: laying hens, cows, chickens and pigs. They then undertook the PC component of the survey followed by the CV.

This order followed Tversky *et al* (1988), and was determined by the aim of each of the methods. The method of paired comparisons precedes the valuations as it is cognitively simpler than that of CV and making a considerable number of paired choices does not translate to a definitive ordering of programmes in the mind of the respondent. Thus, the impact on the valuation scenarios would be minimal, whereas the direct process of revealing preferences within the CVM scenarios may lead to the development and construction of preferences, which could prevent the production of an independent ordering - the main purpose of the paired comparisons exercise. Thus, the respondents choose between pairs of the programmes prior to valuing them.

For notational purposes, all welfare schemes are denoted  $X_j^i$  where X relates to the welfare scheme,  $i$  refers to the method of elicitation (where 1 = PC and 2 = CV) and  $j$  refers to the order of the schemes in order of preference ( $j = 1, 2, 3, 4$ ). Thus, we can state the formal hypothesis that we are testing:

$H_0$ : In the presence of preference reversals, the order of preferences will differ between that obtained from PC and CV. We compare the orderings of the preferences from PC and CV and test the hypothesis  $H_0$ : Order of PC ( $X_1^1, X_2^1, X_3^1, X_4^1$ ) = Order of CV ( $X_1^2, X_2^2, X_3^2, X_4^2$ ).

## 3. Sampling and analytical methodology

The study was administered using face-to face interviews. The sample of 400 respondents was selected from the electoral register of Northern Ireland in a three stage, random sample closely following a design described in Carson *et al.* (1992). The aim of the sampling framework was to ensure that each person within Northern Ireland had an equal probability of being selected, while ensuring that the survey could be practically administered.

<sup>2</sup> To ensure the effectiveness, objectivity and practicality of the welfare improvement schemes, they were based on reports from the Farm Animal Welfare Council (FAWC) and research by the Agricultural Research Institute of Northern Ireland (ARINI).

### Paired Comparisons

In this elicitation method (Edwards, 1957, David, 1988), choose the favourite policy between each of the 6 possible pairs of the 4 alternative policies. From these stated choices, the transitivity of each respondent's choices can be examined, and a numerical scale can be derived for the strength of observed sample preferences for each of the improvement schemes. In order to determine the number of intransitive judgements made by each of the respondents and thus the degree of intransitivity, the formula (3.1) derived by Kendall and Smith (1940) is applied:

$$(3.1) \quad d = \frac{1}{12}n(n-1)(2n-1) - \frac{1}{2} \sum_{j=1}^n a_j^2,$$

Where  $d$  is the number of circular triads (intransitive choices),  $a_j$ , the number of judgments within the set of objects where scheme  $j$  is preferred and  $n$  = number of objects within the set.

Any respondent that has a positive  $d$  value is deemed irrational and excluded from the data set, only transitive judgments reduce into a ranking upon which a scale can be based (Kendall and Smith, 1940). Furthermore, Foster and Mourato (2002) illustrate that even occasional failures of consistency can adversely affect the performance of the standard regression analysis for ranking data. A high level of intransitive respondents would indicate that a rational ordering of schemes is not obtainable.

Data from transitive respondents was then analysed using a discrete choice random utility framework (McFadden, 1974), where the rank is determined on the basis of dominance. This permits a numerical scale is derived based on the probability of a scheme being chosen over another using the model of Bradley and Terry (1952). Assuming that the error terms are distributed logistically, a logit model can be fitted to the observed judgements, enabling estimates of the scale values of each of the welfare schemes to be obtained. The welfare schemes form the parameter of the model, with one scheme set as the origin, thus giving the position of the other schemes relative to it. Thus, the linear index of the logit model for two schemes  $i$  and  $j$ , with  $j$  set as the origin, is specified as:

$$(3.2) \quad Y = \theta_i X_i$$

Where  $Y$  takes the value of 1 for the judgments where object  $i$  is preferred, and 0 when object  $j$  is preferred. This is a simplistic form of the model using only two schemes. To include all the schemes a multivariate logit model is fitted:

$$(3.3) \quad Y = \theta_1 X_1 + \theta_2 X_2 + \dots + \theta_{n-1} X_{n-1}$$

The use of the logit model and its' maximum likelihood estimation permits the statistical testing using t-tests, of the differences between the scale values attributed to the different schemes.

### Contingent valuation

Within the CV study, a double-bounded dichotomous choice format (Haneman, *et al* 1991) was used to obtain estimates of the WTP for each of the welfare improvement schemes, where the payment vehicle was a tax on all food purchases, expressed as an additional amount on the weekly shopping bill.

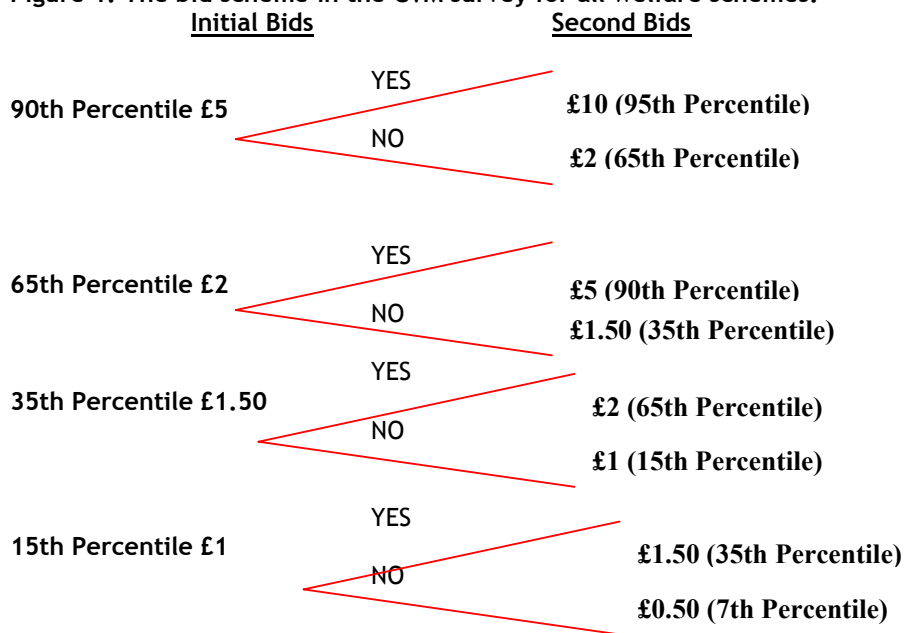
In order to control for possible sequencing effects in valuation (Carson and Mitchell, 1995), an 'exclusive list' framework (Bateman *et al.*, 2004) was used where the schemes are presented as alternatives to any other good given in that list. Here, the expressed value for a good at any position in the list always refers to the same unit of that good,

irrespective of its position, because the reference levels for income, prices, private and public goods across valuation questions remains constant<sup>34</sup>.

The bid amounts given in Figure 1, were determined from an open-ended pilot survey following Hanemann and Kanninen (1999), to ensure that bids were spread across the WTP distribution. This method is considered to ease cognitive effort and it is suitable for goods with a public-good component, such as animal welfare. The stated choices allow one to estimate the WTP distribution.

The responses to the first and second bids where (the value of the second bid is contingent on the response to the first bid (A), if the response was positive a higher bid (A<sub>U</sub>) was offered, whereas if the first bid was rejected, then a lower bid (A<sub>L</sub>) was offered) partitions the WTP distribution as shown below in equation 1 and is illustrated in figure 2.

Figure 1. The bid scheme in the CVM survey for all welfare schemes.



The responses to the first and second bids where (the value of the second bid is contingent on the response to the first bid (A), if the response was positive a higher bid (A<sub>U</sub>) was offered, whereas if the first bid was rejected, then a lower bid (A<sub>L</sub>) was offered) partitions the WTP distribution as shown below in equation 1 and is illustrated in figure 2.

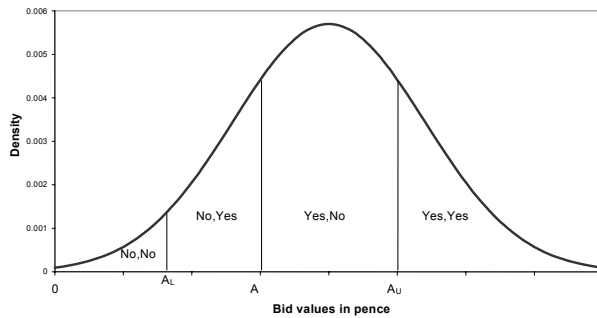
(3.4)

$$\begin{aligned}
 \text{Yes / Yes} &\Rightarrow WTP > A_U \\
 \text{Yes / No} &\Rightarrow A < WTP < A_U \\
 \text{No / Yes} &\Rightarrow A_L < WTP < A \\
 \text{No / No} &\Rightarrow WTP < A_L
 \end{aligned}$$

<sup>3</sup> Any residual variation associated with presentation is therefore a bias (possibly of cognitive psychological origin) and can be termed an *ordering effect*. Empirical evidence for the presence of such effects in CV studies is mixed (Boyle, Welsh and Bishop, 1993).

<sup>4</sup> A split sample bid design was used, to test the value of the good valued last within the sequence. No significant difference was found between the two values.

**Figure 2. Partitioning of the WTP distribution in Double-Bounded Dichotomous Choice CVM.**



In the analysis of the data from the dichotomous choice CVM questions, the observed choices allow one to estimate the underlying WTP distribution with the use of logistic regression. In order to infer estimates of value at the population level to assess the implications of a policy change, a measure of mean WTP was required. Following Duffield and Patterson (1991), a non-negative truncated mean was then calculated from the WTP distribution, with an upper truncation level of 10% ( $C^{+10\%}$ ) to produce conservative estimates of the samples mean WTP, as shown below:

$$(3.5) \quad \tilde{C}^+ = E\{C(q^0, q^1, y, \varepsilon)\} = \int_0^{C^{+10\%}} P(\text{yes}|m, A; \alpha, \beta) dA$$

Where  $q^0$  and  $q^1$  are the ex-ante and post-ante levels of the provision of good  $q$ ,  $y$  the monetary income and  $\varepsilon$  the stochastic elements of the utility function. Confidence intervals of the welfare measures were obtained using the parametric bootstrapping method of Krinsky and Robb (1986) following the approach of Park *et al* (1991).

An ordering of the preferences can be derived from the magnitude of the WTP values, namely, the higher the WTP, the greater the preference. In order to derive a statistically significant preference ordering of the schemes based on the magnitude of mean WTP, the differences between the schemes were tested using Poe *et al.*'s (1994) approach based on convolutions. By applying a Monte Carlo simulation (Krinsky and Robb, 1986) to the differences between the schemes, the significance of those differences was established.

### **Comparing the ordering from PC and CV**

Within this study, the choice-based method of paired comparisons is used in conjunction with the contingent valuation method to examine the consistency of respondent's preferences across choosing and pricing methods. The ordering derived from the choice-based method (PC) is compared with the price-based preference order from the estimated values of the contingent valuation study. This comparison constitutes an external validity test and is conducted to identify any evidence of preference reversals occurring between responses to the two ordering techniques by respondents.

To test the significance of the comparison of the two preference orders in greater depth, a bootstrapping procedure using the Krinsky and Robb (1986) method was conducted to produce a large number of replications of the orderings from both methods to establish how frequently violations of the ordering derived from the methods would occur. Resampling the preference ordering from both CVM and paired comparisons 10,000 times by means of a bootstrapping procedure it was possible to assess the probability of a given ordering occurring in either of the methods and therefore obtain the probability of a preference reversal occurring.

## 4. Results

### *Paired comparisons*

By applying the formula (equation 3.1) of Kendall and Smith (1940), 83% of the sample was found to have entirely transitive preferences. This is similar to the 13% of respondents found to be intransitive by Foster and Mourato (2002), and is low compared to the 46% of respondents found to have inconsistencies in any of the tests for transitivity, rank consistency or dominance.

A logistic regression (equation 3.3) was applied to the judgements of the transitive respondents<sup>5</sup> following the Bradley and Terry model (1952), with the resulting parameter estimates of the scale values presented in Table 4.1. Based on the scale values, the order of preferences for the improvement schemes was found to be chickens, followed by laying hens, dairy cows and lastly pigs.

**Table 4.1 Estimates of scale values from paired comparisons data.**

Scheme	Scale Values
Chickens	0.7921 (0.0816)
Laying Hens	0.7586 (0.0819)
Dairy Cows	0.6334 (0.0813)
Pigs	0 (0)

The above order of preferences does not take into account the variability of the estimated scale values, thus requiring an assessment of the statistical significance of the differences between the scale values of each of the schemes. This is undertaken using *t*-tests, with the result shown in Table 4.2.

**Table 4.2. Test of Ho: No significant differences of the preference scale values from paired comparisons.**

Laying Hens Vs Chickens	Accept
Laying Hens Vs Dairy Cows	Accept
Chickens Vs Dairy Cows	Accept
Laying Hens Vs Pigs	Reject***
Chickens Vs Pigs	Reject***
Dairy Cows Vs Pigs	Reject***

\* 10% confidence level, \*\*5% confidence level, \*\*\* 1% confidence level

A significant difference was only observed between the scale value of the scheme to improve the welfare of the pigs compared to each of the other schemes. In all such comparisons, respondents chose to improve the welfare of all the other species in preference to the pig welfare improvement scheme. No significant differences were found between the scale values for the schemes to improve the welfare of laying hens, dairy cows and chickens. Thus the order of preferences from the paired comparisons study is that the pig welfare improvement scheme is least preferred, while respondents express equal preference for the schemes to improve the welfare of the laying hens, chickens and dairy cows.

<sup>5</sup> The preferences of 17% of the sample who had intransitive respondents were excluded from further analysis (Kendall and Smith, 1940; Foster and Mourato, 2002).

### Contingent Valuation

We present a parsimonious specification of the estimated non-negative truncated (10%) mean which included the explanatory variable of household income. Each of the parameters within the model for the relevant variables, bid amount and household income had the expected signs and were statistically significant as shown in table 4.3, with the means and 95% confidence intervals of the welfare measures are presented in Table 4.4, together with the aggregate estimates for annual total economic value.

**Table 4.3. Estimates for the double-bounded CVM models**

	Parameter	Coefficients	Mean WTP ( $C^{+ \sim 10\%}$ )
<b>Laying Hens</b> (N=192) Log-Like= -249.48	$\alpha$	1.29	£2.95
	(Standard Error)	(0.2162)	
	$\beta$	-0.585	
	(Standard Error)	(0.0517)	
	Household Income	1.217	
	(Standard Error)	(0.3084)	
<b>Chickens</b> (N=192) Log-Like= -235.94	$\alpha$	1.375	£2.63
	(Standard Error)	(0.2159)	
	$\beta$	-0.682	
	(Standard Error)	(0.0552)	
	Household Income	1.196	
	(Standard Error)	(0.3158)	
<b>Dairy Cows</b> (N=191) Log-Like= -255.3	$\alpha$	1.394	£2.89
	(Standard Error)	(0.225)	
	$\beta$	-0.603	
	(Standard Error)	(0.0508)	
	Household Income	0.936	
	(Standard Error)	(0.3312)	
<b>Pigs</b> (N=192) Log-Like= - 243.62	$\alpha$	1.301	£2.10
	(Standard Error)	(0.2127)	
	$\beta$	-0.791	
	(Standard Error)	(0.0642)	
	Household Income	0.921	
	(Standard Error)	(0.318)	

**Table 4.4. Estimated Mean WTP for welfare schemes from double-bounded CVM**

	Individual Weekly WTP (95% c.i.)	Annual Total Economic Value (95% c.i.)
<b>Laying Hens</b>	£2.95 (2.36-3.31)	£81.3 Million (73.3 - 89.6 million)
<b>Dairy Cows</b>	£2.89 (2.30-3.24)	£72.5 Million (65.6 - 79.7 million)
<b>Broiler Chickens</b>	£2.63 (2.13-2.96)	£79.7 Million (71.7 - 87.6 million)
<b>Pigs</b>	£2.10 (1.62-2.36)	£57.9 Million (52 - 87.6 million)

An order of preferences for the welfare improvement schemes can be gained through a comparison of the point estimates displayed, however this does not take into account the variability of the estimates. Thus, statistical testing of the point estimates was undertaken using the bootstrapping method (Krinsky and Robb, 1986).

**Table 4.5. Tests of Ho: No significance differences in mean WTP between schemes**

Laying Hens Vs Chickens	Accept
Laying Hens Vs Dairy Cows	Accept
Chickens Vs Dairy Cows	Accept
Laying Hens Vs Pigs	Reject***
Chickens Vs Pigs	Reject***
Dairy Cows Vs Pigs	Reject***

\* 10% confidence level, \*\*5% confidence level, \*\*\* 1% confidence level

As can be seen in Table 4.5, a significant difference was observed between the WTP for the pig welfares improvement scheme compared to each of the other schemes. In all such comparisons, respondents were WTP less for pig welfare than for the welfare of the other species. No significant differences were found between the WTP for schemes to improve the welfare of laying hens, dairy cows and chickens. Thus the order of preferences from the contingent valuation study is that the pig welfare improvement scheme is least preferred, while respondents express equal preference for the schemes to improve the welfare of the laying hens, chickens and dairy cows.

#### *Testing for the Presence of Preference reversals*

Both methods of preference elicitation independently produced the same statistically significant order of preference, with the welfare improvement scheme for pigs being the least preferred of all the schemes and indifference expressed between the improvement schemes for laying hens, chickens and dairy cows. Therefore it would appear that -- in this case study - no significant preference reversal has occurred between elicitation methods, as the same preference ordering has been found by the cost-related contingent valuation and the rank-related paired-comparison method.

The significance of this finding was tested through a bootstrapping procedure described in section 3. Due to the indifference between the schemes for laying hens, chickens and dairy cows, we hypothesise that a change in ordering can only occur when the scheme to improve the welfare of pigs is ranked anywhere other than least preferred by either method. Thus, in testing for preference reversals between elicitation methods, differences in orders of the schemes to improve the welfare of laying hens, chickens and dairy cows can be disregarded, with account only taken of whether the scheme to improve the welfare of pigs is always least preferred or placed above any of the other improvement schemes.

Resampling the preference ordering from both CVM and paired comparisons 10,000 times by means of a bootstrapping procedure it was possible to assess the probability of the pig welfare scheme being placed anywhere other than below the other schemes in either of the elicitation methods and for potential preference reversals - where the preference ordering differs between methods.

**Table 4.6 Probability of the specified preference ordering occurring as estimated for paired comparisons and contingent valuation methods**

Method	No. of occurrence of pigs NOT been the least preferred scheme out of 10,000 replications	Probability of specified order occurring
Paired Comparisons	0	100%
Contingent Valuation Model:	120	98.8%

In this study, as can be seen in Table 4.6, in 10,000 replications, pigs were always least preferred in the observed ordering derived from the paired comparisons data. From the contingent valuation estimates, the scheme for pigs was not least preferred only in 120 of the 10,000 recorded simulations. This shows that from the CVM data there is approximately 1.2% probability of a different order of preferences than the pig scheme being least preferred being produced. Thus, there is a 1.2% probability of a preference reversal

occurring between the contingent valuation and paired comparisons methods in this case study.

## 5 conclusions

The findings of this study concur with previous studies outlined previously by suggesting that substantial public support exists for improving the welfare of farm animals within the UK. Furthermore, respondents appear to be largely rational in their decisions, with a low number of intransitive responses reported and no significant preference reversals observed between elicitation methods at the 98% confidence level. There is a clear preference for any of the investigated schemes for laying hens, chickens and dairy cows to be implemented over the pig improvement scheme, with indifference expressed between the three preferred schemes.

With little evidence of preference reversals being found between the contingent valuation and the paired comparisons surveys, and low levels of intransitivity, it can be concluded that on average respondent choices were rational with stable, well-formed preferences between the proposed schemes to improve the welfare of pigs and the other three programmes. The lack of preference reversals verifies the use of both these methods of preference elicitation.

The concern raised by Loomes (1999) that the elicitation method may have an influence on the preferences and values of the public and the consequent fear that the use of CVM may be potentially dangerous in policy decisions are not supported in this study. Thus, the estimates of value obtained can be considered to be valid and applicable for use within decisions regarding future policy changes relating to farm animal welfare, with the caveat that the estimated values of WTP relate directly to policy changes based on these specific welfare improvement schemes. It would therefore appear that preference reversals found by economic psychologists and experimentalists in risky choices are not found in this field study involving technically complex and emotive choices on welfare improvement.

While respondents have consistently reversed their preferences, while seemingly unaware that they are doing so, the situations in which all the experiments carried out by the psychologists and experimental economists are highly unrealistic, and this lack of realism may have caused the preference reversals rather than persistent violations of procedural invariance. For example, the study by Irwin *et al* (1993) shows preference reversals occurring in comparisons between the market commodities and environmental amenity with no reversals for comparisons involving the two market commodities. A simple explanation for these observed preference reversals is the degree of experience possessed by the subjects for the each of the goods in question.

Cummings *et al* (1986) conjectured that the reliability of estimated WTP would be less for public good than for private goods, as public goods lack market prices and people have limited experience of valuing them. Peterson and Brown (1998) from a paired comparisons study involving public and private goods together with sums of money, observed significant differences in the time taken to make the different types of comparison. Choices involving monetary amounts took the shortest time, followed by comparisons between private goods with the longest times for choices between public goods. They suggested that this finding is consistent with the possibility that respondents convert each good into a monetary amount before they make a choice. As previously stated, there is an implicit assumption made that in judging the subject "*must mentally construct some function of the relevant characteristics and use this as a basis of comparison*" (David, 1988, p4). With this function being based on a monetary amount, this conversion will be easier for a private good, where the subject will be more familiar with the relevant monetary values, than for a public good. Thus, comparing a public good either with other public goods or with a private good, for both pricing and choosing decisions is expected to be difficult, relative to choosing between two private goods.

This is reflected in an assessment undertaken by Brown *et al* (1999), of inconsistent choices made between all types of comparisons. There are significantly lower numbers of inconsistent choices made for comparisons involving monetary amounts, with the number of inconsistent choices made over a public and a private good being significantly greater than choices between two private goods.

The implication of the findings of the studies by Peterson and Brown (1998) and Brown *et al* (1999), is that the preference reversals observed by Irwin *et al* (1993) may simply be the result of respondents finding it more difficult to choose between a public and a private good. Thus, in making decisions over non-like goods, the complex cognitive processes required can result in intransitivity, within a single method and/or preference reversal when more than one method is used.

Furthermore, while it is possible that preference reversals have been manufactured through the unrealistic nature of both the objects and the procedures used to elicit preferences, there also exists the possibility that they have resulted from poor statistical analysis of the data.

Within all of the studies (Irwin *et al* 1993, Brown 1984), while differences in preferences have been observed to occur between those produced from the contingent valuation method and the other choice method used, no statistical testing has been conducted to ensure whether these differences are actually significant. Furthermore, with the open-ended format elicitation method being used throughout these two CVM studies, the validity of the derived estimates of WTP is suspect as this format is overly sensitive to the hypothetical CV scenario and prone to strategic bias. The non-incentive compatible nature of this valuation task is apparent as it would be rational for a respondent suspecting that they might be asked afterwards to purchase the goods in question to place a lower price on their preferred option. The mechanism is clearly open to this type of strategic response. This effect is exacerbated within the Irwin *et al* (1993) study as the preferences for each of the objects from the valuation study is derived through ordering the dollar amounts offered for each object.

### **Conclusions**

We have argued that a multifunctional agricultural system can provide society with a valuable public good: increased farm animal welfare. This is an issue that lends itself to divergence of opinion mostly because of its emotional undertones and the diversity of people's views. Nevertheless policy making requires guidance as to the intensity and the relative magnitudes of benefits associated with legislative efforts to address this very current issue. In this paper we attempt to assess the consistency of two stated preference methods in capturing both the intensity and the ordering of preferences for four different farm animal welfare enhancing schemes.

We conclude that members of the Public make largely rational stated choices over alternative farm animal welfare enhancing schemes. Similar orderings are obtained by the classic contingent valuation and the less conventional paired comparison preference elicitation schemes. In this field study these methods appear largely free of the most radical reversal and framing effects raised by economic psychologists e.g. Tvesky *et al* 1990.

### **References**

- Adamowicz W, Louviere J and Williams M (1994) Combining revealed and stated preference methods for valuing environmental amenities *Journal of Environmental Economics and Management*, 26 (3), 271-292.
- Bateman, I.J., Cole, M., Cooper, P., Georgiou, S., Hadley, D. and Poe, G.L., (2004) On visible choice sets and scope sensitivity, *Journal of Environmental Economics and Management*

- Bennett, R.M. and Blaney R. (2003) Estimating the benefits of farm animal welfare legislation using the contingent valuation method. *Agricultural Economics* 29: 85-98
- Bennett, R.M. and Larson, D. (1996) Contingent Valuation of the Perceived Benefits of Farm Animal Welfare Legislation: An Exploratory Survey. *Journal of Agricultural Economics*, 47(2): 224-235.
- Berg, J., Dickhaut, J. and O'Brien, J. (1985), Preference Reversal and Arbitrage in *Research in Experimental Economics* (Vol. 3), V. Smith (Ed), Greenwich, Conn.: JAI Press, 31-72.
- Blandford, D. and Fulponi, L. (1999) Emerging public concerns in agriculture: domestic policies and agricultural trade commitments, *European Review of Agricultural Economics*, 26(3):409-424.
- Boyle, K.J., Welsh, M.P., and Bishop, R.C. (1993) The Role of Question Order and Respondent Experience in Contingent Valuation Studies *Journal of Environmental Economics and Management* 25: S80-S99.
- Bradley, R.A. and Terry, M.E. (1952) Rank Analysis of Incomplete Block Designs. I. The Method of Paired Comparisons. *Biometrika*, 39: 324-45.
- Brown, T.C. (1984), The Concept of Value in Resource Allocation *Land Economics* 60: 231-46.
- Brown, T.C., Peterson, G.L., Clarke, A. and Birjulin, A. (1999), Reliability of Individual Valuations of Public and Private Goods Paper Presented at the ninth Annual Conference of the European Association of Environmental and Resource Economists held at the University of Oslo, Norway, 25-27<sup>th</sup> of June 1999.
- Camerer, C. (1995) Individual Decision Making. In Kagel, J.H. and Roth, A.E. (eds). *The Handbook of Experimental Economics*. Princeton University Press: Princeton.
- Carson, R.T. and Mitchell, R.C. (1995), Sequencing and Nesting in Contingent Valuation Surveys, *Journal of Environmental Economics and Management* 28: 155-173.
- Carson, R.T., Groves, T. and Machina, M. (1999) Incentive and Informational Properties of Preferences Questions. Plenary Address, European Association of Environmental and resource Economists, Oslo, Norway, June 1999.
- Carson, R.T., Mitchell, R.C., Hanemann, W.M., Kopp, R.J., Pressor, S. and Rudd, P.A. (1992) A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill. Report to the Attorney General of the State of Alaska, November 1992.
- Chu, Y.-P. and Chu, R.-L. (1990) The subsistence of preference reversals in simplified and marketlike experimental settings. *American Economic Review*, 80: 902-911.
- Costello, C.A. (2000) *Minimising Design Bias in a Contingent Valuation Study Of Farm Animal Welfare Improvements*. Ph.D.Thesis, Department of Agricultural and Food Economics, The Queen's University of Belfast.
- Cox, J. and Grether, D. (1996) The Preference Reversal Phenomenon: Response Mode, Markets and Incentives. *Economic Theory*, 7(3): 381-405.
- Cummings, R.G., Brookshire, D.S. and Schulze, W.D. (1986), *Valuing Environmental Goods - An Assessment of the Contingent Valuation Method*, Totowa: Rowman & Allenheld.
- David, H.A. (1988) *The method of paired comparisons 2<sup>nd</sup> Edition*. Charles Griffin & Company Ltd: UK.
- Duffield, J.W. and Patterson, D.A. (1991) Inference and Optimal Design for a Welfare Measure in Dichotomous Choice Contingent Valuation. *Land Economics*, 67(2): 225-39.
- Edwards, A.L. (1957) *Techniques of attitude scale construction*. Appleton-Century-Crofts, Inc: New York.
- Foster V. and Mourato, S. (2002) Testing for Consistency in Contingent Ranking Experiments. *Journal of Environmental Economics and Management*, 44: 309 - 328.
- Frykblom, P. and Andersson, J. (2002) Exploring non-market values for the social impact of farm animal welfare legislation. Paper presented at the 2<sup>nd</sup> World Congress of Environmental and Resource Economists, Monterey, USA, June 2002.
- Glass, C.A.; Hutchinson, W.G. Beattie, V.E. (2001) Investigating issues of overstatement in a CVM study of WTP for pig welfare improvements. *Paper presented at the AES Conference, Harper-Adams University College September 2001*.
- Grether, D. and Plott, C. (1979) Economic Theory of Choice and the preference Reversal Phenomenon *American Economic Review* 69: 623- 638.

- Hanemann, W. M., Loomis, J. and Kanninen, B.J. (1991), Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation, *American Journal of Agricultural Economics* 73(4): 1255-1263.
- Hanemann, W.M. and Kanninen, B. (1999) The Statistical Analysis of Discrete-Response CV Data. In: Bateman, I.J. and Willis, K.G. *Valuing Environmental Preferences - Theory and Practice of the CVM in the US, EU, and Developing Countries*. Oxford University Press: Oxford.
- Irwin, J.R., Slovic, P., Lichtenstein, S. and McClelland, G.A. (1993) Preference Reversals and the Measurement of Environmental Values. *Journal of Risk and Uncertainty*, 6: 5-18.
- Kahneman, D. and Tversky, A. (1979), Prospect Theory: An Analysis of Decision Under Risk *Econometrica* 47(2): 263-291.
- Kendall, M.G. & Smith, B.B. (1940) On the Method of Paired Comparisons. *Biometrika*, 31: 324-45.
- Krinsky I. and Robb, A.L. (1986) On Approximating the Statistical Properties of Elasticities. *Review of Economics and Statistics*, 68: 715-19.
- Lichtenstein, S. and Slovic, P. (1971) Reversals of Preference between Bids and Choices in Gambling Decisions. *Journal of Experimental Psychology*, 89(1): 46-55.
- Lindman, H.R. (1971) Inconsistent Preferences Among Gambles. *Journal Of Experimental Psychology*, 89(2): 390-397.
- Loomes, G. (1998) Probabilities Vs Money: A Test of Some Fundamental assumptions About Rational Decision Making. *The Economic Journal*, 108: 477-489.
- Loomes, G. (1999) Some Lessons from Past Experiments and Some Challenges for the Future. *The Economic Journal*, 109: F35-F45.
- McFadden, D. (1974) Conditional Logit Analysis of Qualitative Choice Behaviour. In P. Zarembka, ed., *Frontiers in Econometrics*, Academic Press: New York, 105-142.
- Mitchell, R. C. and Carson, R.T. (1989), Using Surveys to Value Public Goods: The Contingent Valuation Method, Washington D.C.: Resources for the Future.
- Park, T., Loomis, J.B., and Creel, M. (1991) Confidence Intervals for Evaluating Benefits Estimates from Dichotomous Choice Contingent Valuation Studies. *Land Economics*, 67(1): 64-73.
- Peterson, G.L. and Brown, T.C. (1998), Economic Valuation by the Method of Paired Comparison, With Emphasis on Evaluation of the Transitivity Axiom. *Land Economics* 74(2): 240-61.
- Poe, G.L., Severance-Lossin, E.K. and Welsh, M.P. (1994) Measuring the Difference (X-Y) of Simulated Distributions: A Convolutions Approach. *American Journal of Agricultural Economics*, 76: 904-915.
- Pommerehne, W.W., Schneider, F. and Zweifel, P. (1982), Economic Theory of Choice and the Preference Reversal Phenomenon: A Re-examination *The American Economic Review* 72(3): 569-574.
- Reilly, R.J. (1982), Preference Reversal: Further Evidence and some Suggested Modifications in Experimental Design *The American Economic Review* 72(3): 576-584.
- Rolfe, J.C. (1999) Ethical Rules and the Demand for Free Range Eggs, *Economic Analysis and Policy*, 29(2):187-206.
- Tversky A., Sattath, S. and Slovic, P. (1988b) Contingent Weighting in Judgment and Choice. *Psychological Review*, 95(3): 371-384.
- Tversky, A., Slovic, P. and Kahneman D. (1988a) The determinants of preference reversals. Unpublished
- Tversky, A. (1969) Intransitivity of Preferences. *Psychology Review*, 76(1): 31-48.
- Tversky, A., Slovic, P. and Kahneman D. (1990) The Causes of Preference Reversal. *The American Economic Review*, 80(1): 204-217.