University of Kent

School of Economics Discussion Papers

A General Treatment of Non-Response Data From Choice Experiments Using Logit Models.

Kelvin Balcombe and Iain Fraser

October 2009

KDPE 09/16

A GENERAL TREATMENT OF NON-RESPONSE DATA FROM CHOICE EXPERIMENTS USING LOGIT MODELS.

Kelvin Balcombe
University of Reading

Iain Fraser
University of Kent

ABSTRACT:

A new approach is developed for the treatment of 'Don't Know'(DK) responses, within Choice Experiments. A DK option is motivated by the need to allow respondents the opportunity to express uncertainty. Our model explains a DK using an entropy measure of the similarity between options given to respondents within the Choice Experiment. We illustrate our model by appling it to a Choice Experiment examining consumer preferences for nutrient contents in food. We find that similarity between options in a given choice set does explain the tendency for respondents to report DK.

KEYWORDS: Choice Experiment, Respondent Uncertainty, Bayesian Methods

JEL CLASSIFICATION: C35, I18, Q18.

Acknowledgements: We gratefully acknowledge the financial support provided by the Food Standards Agency in supporting the Choice Experiment that provides the data used in this paper.

Addres for Correspondence: Iain Fraser, School of Economics, University of Kent, Canterbury, Kent, CT2 7NP, Phone: +44 (0)1227823512. Email: i.m.fraser@kent.ac.uk

A General Treatment of Non-Response Data from Choice Experiments using Logit Models.

1 Introduction

Within a Choice Experiment (CE) many studies will include either a Status Quo option or an opt-out option. The inclusion of either of these options in a CE, it is argued, ensure that the choices being made are such that consistent welfare estimates will be derived (eg. Hensher et al., 2005). To date a number of papers have considered the form that these options might take in a CE and how this impacts the results derived, e.g., Banzhaf et al. (2001), Kontoleon and Yabe (2003) and Fenichel et al. (2009). Surprisingly, and unlike the Contingent Valuation (CV) literature, CEs are generally not designed with a Don't Know (DK) option so as to capture respondent uncertainty. An exception is Hatton MacDonald et al. (2005) who include a Don't Know option in their CE although no explanation for its inclusion or how the data are used is provided.

The inclusion of the DK option in stated preference survey instruments can be traced back to the recommendation of the NOAA Panel and the inclusion of the "No-Answer" option (Haab and McConnell, 2000) in CV studies. Within the CV literature the inclusion of the DK option has lead to the development of various approaches to explain the reasons for a DK response. Examples in the literature include Welsh and Poe (1998), Evans et al. (2003), and Balcombe and Fraser (2009). The motivation for the various approaches found in the literature are that DK responses (or more generally uncertain responses) may be informative and such needs to be included within model estimation.

To date there is only a very limited literature on how uncertainty might impact CE. Fenichel et al. (2009) examine the possibility that in some choice situations respondents are either uncertain or indifferent to some of the options presented, and for this reason are unable to respond in a definitive way. To examine this issue they employ four alternative response formats within a single CE. They found that employing different forms of the opt-out option, as well as including two options in a single choice set, that this significantly impacted their resulting willingness to accept estimates. Of most significance is the finding that by allowing two opt-out options in the same choice set they cannot only capture respondent uncertainty but also indifference with respect to the choices on offer. Both Lundhede et al. (2009) and Kosenius (2009) takes a different approach. The CEs employed in these papers follows many of the CV studies and employs uncertainty scales after the choices have been made. Using the estimated utility function in combination with the uncertainty data they find that response certainty increases with an increase in the difference in the utility derived from the alternative choices within a given choice set. This indicates that CEs which are designed employing utility balance (ie, Huber and Zwerina, 1996) are more likely to yield uncertain responses.

In this paper we take a different approach to examine uncertainty in a CE. Specifically, we employ a CE that includes both a Status Quo and a DK option. We then develop a treatment of the DK responses in a CE using the Conditional Logit model that allows us to examine the underlying motivation of DK responses directly. The approach that we take draws on the method to treat DKs developed by Balcombe and Fraser (2009) for Dichotomous Choice CV. Essentially, probabilities are assigned to the event that a respondent who has the highest utility for each option within a choice set, instead reports DK, and conversely a probability that a respondent replying DK actually derives the highest utility from a given option. Our model postulates that these probabilities will depend on the similarity between some or all of the options presented to them (within a given choice set). By 'similarity' we mean that the utility derived from two options are predicted to be close.

Within the estimation of the Conditional Logit, the probability that an individual will choose a given option is calculated. This probability is then used to calculate an entropy measure estimating the similarity between a given set of choices within a given choice set. We employ the Shannon measure of entropy which has been employed in a number of different econometric contexts (see Golan et al., 1996, for details). If the similarity measure can be used to predict the frequency of DK responses, this represents evidence that DKs are informative and since the preference parameters are used to derive the measure of similarity, their inclusion aids the estimation of these parameters. If the entropy measure cannot explain whether a DK occurs, this indicates that there are other reasons why DKs are occurring (e.g. a set of individuals simply do not understand the CE or cannot be bothered to respond in a meaningful way). This can be dealt with using simple restrictions to the general model that we develop. Thus, unlike the existing literature our modified Conditional Logit model offers a general approach to testing the potential impact of respondent uncertainty.

To implement this new model we derive the likelihood functions for a new class of Logit models that are able to deal with DK responses in the manner described. These models could be estimated using classical or Bayesian procedures. However, we employ a Bayesian Monte Carlo Markov Chain (MCMC) approach to estimation as this provides a convenient and general framework for model comparison through the calculation of the marginal likelihood. For a recent application of Bayesian procedures using MCMC to estimate CE data see Balcombe et al. (2009).

To demonstrate the utility of the model developed we examine a CE designed to understand how consumers respond to the UK food label Traffic Light System (TLS). The TLS has been recommended by the UK Food Standards Agency (FSA) as an industry wide approach to nutritional labelling. The TLS indicates if the main nutrients in food (i.e., Salt, Sugar, Fat and Saturated Fat) are high (Red), medium (Amber) and low (Green). The system is relatively simple. A Red light indicates a (excessive) high level of a specific nutrient, Amber is medium and Green low. The mix of nutrient colours on a food item is derived from the quantity of each of the nutrients per 100 grams of the food item. Thus, the system is simple and it implies that by selecting foods with Green and maybe some Amber lights this choice can be considered a healthy option. Note, our CE adds to small but growing number of studies that examine food labels and health such as Berning et al. (2008) and Gao and Schroeder (2009).

Our key finding is that people do seem to reply DK to those choice sets which are estimated to have similar utility. In particular, we find that when the top two preferred options present the respondent with similar utility, respondents are most likely to reply DK as opposed to the circumstance where all three options provide a similar level of utility. This results raises questions regarding the use of efficient utility balance designs in CE if respondent uncertainty is to be reduced. On a policy level our results indicate that consumers appear to behave in a manner consistent with our expectations regarding the impact of the TLS food label. In particular, we identify a strong preference on the part of respondents to avoid a basket of goods containing a mix of foods with any Red lights.

The paper proceeds in second section by outlining the model and particular parameterization that we use for the model along with the priors that we employ. The third section introduces the empirical results and the last section concludes.

2 Model specification

Let x'_{ij} denote a vector or attributes, where j denotes the jth option and i the ith choice situation. Utility (u_{ij}) derived from x'_{ij} is a function of the following form:

$$
\begin{aligned}
&(1) \qquad & u_{ij} &= x'_{ij}\beta + e_{ij} \\
&(2) \qquad & i &= 1, ...n \\
&(3) \qquad & j &= 1, ...J
\end{aligned}
$$

where the errors e_{ij} are Gumbel distributed and are assumed to identically an independently distributed across i and j. The probability that the basket x'_{ij} is preferred therefore has a logistic distribution of the form:

$$(4) \qquad P\left(u_{i,j} = Max\left(u_{i,1}...., u_{i,k}\right)\right) = p_{ij} = \frac{e^{x'_{ij}\beta}}{\sum e^{x'_{ij}\beta}}$$

In the theory that follows, the preferred option may not be reported (since a DK might be reported). Consequently, define:

$$(5) \qquad \delta_{ij} = 1 \text{ if } u_{i,j} = Max\left(u_{i,1}....u_{i,k}\right)$$

It follows that (as defined in (4)) :

$$(6) \qquad P\left(\delta_{ij} = 1\right) = p_{ij}$$

Within this approach we do not observe δ_{ij} but another variable y_{ij}.

$$
\begin{aligned}
y_{ij} &= 1 \text{ if } j\text{th option is reported as the preferred option in choice situation } i \\
&= 0 \text{ otherwise.}
\end{aligned}
$$

However, if the respondent reports a DK, then $\sum_{j=1}^{J} y_{ij} = 0$. Therefore, further define:

$$
\begin{aligned}
(7) \qquad \epsilon_i &= I\left(\sum_{j=1}^{J} y_{ij} = 1\right) = 1 \text{ if preference reported} \\
&= 0 \text{ otherwise (i.e. a DK is reported)}
\end{aligned}
$$

where I represents the information content. For the purposes of exposition, we shall first assume that the propensity to report DK does not vary across i. We will then generalize this approach to allow these probabilities to depend on the similarity of choices within the ith choice circumstance.

First, we define the following parameters:

$$\begin{aligned} (8)\quad \theta_{k|j} &= \text{probability reporting option } k \text{ given preference for option } j \\ \theta_{\bullet|j} &= \text{probability reporting DK given preference for option } j \end{aligned}$$

In turn it follows that:

$$\text{if } \theta_{\bullet|j} + \theta_{j|j} = 1 \text{then } \theta_{k|j} = 0 \text{ if } k \neq j \tag{9}$$

Conversely, we can define the probability that the jth option is preferred, given that the respondent has reported a DK as:

$$\psi_{j|\bullet} = P\left(\delta_{ij} = 1 | \sum_{j=1}^{J} y_{ij} = 0\right) \tag{10}$$

Axiomatically, it follows that (from quantities defined in (8) and (4)):

$$\begin{aligned} \psi_{j|\bullet} &= \frac{P\left(\epsilon_i = 0|\delta_{ij} = 1\right) P\left(\delta_{ij} = 1\right)}{\sum_{j=1}^{J} P\left(\epsilon_i = 0|\delta_{ij} = 1\right) P\left(\delta_{ij} = 1\right)} \\ &= \frac{\theta_{\bullet|j} p_{ij}}{\sum_{j=1}^{J} \theta_{\bullet|j} p_{ij}} \end{aligned} \tag{11}$$

Let y_i be the ith set of responses, Y be the set of all responses from all respondents under all choice circumstances, and $f(Y|\beta,\Theta)$ ($f(y_i|\beta,\Theta)$) denote the conditional density of Y (y_i) given the parameters in the utility function β along with the parameters $\Theta = \left\{\theta_{\bullet|j}\right\}_j$. This represents the likelihood function:

$$f(Y|\beta,\Theta) = \prod_{i=1}^{n} f(y_i|\beta,\Theta) \tag{12}$$

It follows that:

$$
\begin{aligned}
f\left(y_{ij}=1|\beta,\Theta\right) &= \sum_{k=1}^{J} f\left(y_{i,j}=1|\delta_{i,k}=1,\beta,\Theta\right)p_{ik} \\
&= \left(1-\theta_{\bullet|j}\right)p_{ij} \qquad (13)
\end{aligned}
$$

and:

$$
\begin{aligned}
f\left(\epsilon_i=0|\beta,\Theta\right) &= \sum_{k} f\left(y_{i,j}=0|\delta_{i,k}=1,\beta,\Theta\right)p_{ik} \qquad (14) \\
&= \sum_{k} \theta_{\bullet|k}p_{ik}
\end{aligned}
$$

Thus, the likelihood function is:

$$
f\left(Y|\beta,\Theta\right) = \prod_{i=1}^{n}\left(\sum_{k=1}^{J}\theta_{\bullet|k}p_{ik}\right)^{1-\epsilon_i}\prod_{j=1}^{J}\left(1-\theta_{\bullet|j}\right)^{y_{ij}}\prod_{j=1}^{J}p_{ij}^{y_{ij}} \qquad (15)
$$

2.1 Measuring Similarity in Utility

We can generalize the approach described above by introducing an entropy measure into the logistic model upon which the probability of choosing DK depends. Our approach adds is in the spirit of a growing number of papers that examine how the degree of attribute variation within choice sets can impact the respones made. For example, Dellaert et al. (1999) and DeShazo and Fermo (2002) both model choice consistency employing a heteroskedastic multinomial logit specification.

Specifically, we employ the Shannon measure of entropy were for a random variable x with K possible outcomes $(k = 1, 2, ..., K)$ and associated probabilities p_k the entropy of the probabilities $p = (p_1, p_2, ...p_K)$‘ is

$$
H(\mathbf{p}) = -\sum_{k=1}^{K} p_k \ln p_k \qquad (16)
$$

where 0 times $\ln(0) = 0$. Thus, the measure of entropy H, which is normally taken to be a measure of uncertainty for the K events will be maximised (or the positive sum minimised) when the probabilities are equal.

That is, we have a uniform distribution (Golan et al, 1996, page 8). We introduce this measure into the logistic as follows:

$$\theta_{\bullet|k,i} = \frac{\exp\left(\alpha_{0,k} + \alpha_{1,k} \sum_{j=1}^{J} p_{ij} \ln\left(p_{ij}\right)\right)}{1 + \exp\left(\alpha_{0,k} + \alpha_{1,k} \sum_{j=1}^{J} p_{ij} \ln\left(p_{ij}\right)\right)} \tag{17}$$

In (17) the entropy measure is included in both numerator and denominator.

While general, the formulation in (17) does not, without further restrictions, have a strong motivation unless the ordering of the options within the choice set have a common structure throughout the choice set. For example, if the options are always labelled in a way that is consistent with their ordering, respondents response to 'similarity' may depend on which option they prefer.

In this paper we consider three specific models:

Model 1: In model one we assume that $\alpha_1 = 0$, in which case

$$\theta_i = \frac{\exp\left(\alpha_0\right)}{1 + \exp\left(\alpha_0\right)} \tag{18}$$

In (18) the probability of reporting DK, is simply a constant, and does not depend on whether the options are similar.

Model 2: In model two we introduce the entropy measure and assume that $\alpha_{0,k} = \alpha_0$ and $\alpha_{1,k} = \alpha_1$ for all k such that we obtain

$$\theta_i = \frac{\exp\left(\alpha_0 + \alpha_1 \sum_{j=1}^{k} p_{ij} \ln\left(p_{ij}\right)\right)}{1 + \exp\left(\alpha_0 + \alpha_1 \sum_{j=1}^{k} p_{ij} \ln\left(p_{ij}\right)\right)} \tag{19}$$

This is the most general form of the model modified logit model in that it includes the entropy measure of similarity between all of the options within a given choice set. Thus, if there are four options in a choice set k would be equal to four.

Model 3: In model three it is not the overall similarity between all choices, but the top two choices that determines whether the respondent replies DK. That is, if they can discount options with lower utility they are left with two options that are relatively similar in terms of their utility. In this case we have

$$\theta_i = \frac{\exp\left(\alpha_0 + \alpha_1 \sum_{j=1}^{2} p_{ij}^* \ln\left(p_{ij}^*\right)\right)}{1+\exp\left(\alpha_0 + \alpha_1 \sum_{j=1}^{2} p_{ij}^* \ln\left(p_{ij}^*\right)\right)} \tag{20}$$

where p_{i1}^* is largest element of $(p_{i1,}..p_{ik})$ and p_{i2}^* is the next highest element. Thus, model 3 allows the researcher to examine pairwise choices within

a larger set. This is an interesting model to test especially if a utility balance approach has been employed as part of the experimental design.

Accordingly, the likelihood function for this modified logistic model now becomes a function of the parameters $\beta, \alpha = \{\alpha_i\}$

$$f\left(Y|\beta, \alpha\right), \tag{21}$$

whereby, the likelihood simplifies to:

$$f\left(Y|\beta, \alpha\right) = \prod_{i=1}^{n} \theta_i^{1-\epsilon_i} \prod_{j=1}^{J} (1-\theta_i)^{y_{ij}} \prod_{j=1}^{J} p_{ij}^{y_{ij}}. \tag{22}$$

The likelihood in (22) above can be estimated using this likelihood function under classical procedures. Alternatively (as in this paper) prior distributions $f\left(\beta, \alpha\right)$ could be assigned to the parameters and the posterior could be derived using the proportionality:

$$f\left(\beta, \alpha|Y\right) \propto f\left(Y|\beta, \alpha\right) f\left(\beta, \alpha\right). \tag{23}$$

The priors we employ in this study are independent multivariate normal.

$$\Theta = \left(\beta', \alpha'\right)' \tag{24}$$

and

$$\Theta \sim N\left(0, V_0\right), \tag{25}$$

where V_0 denotes a variance matrix with all off diagonal elements equal to zero.

2.2 Reparameterization into WTP space

Calculating the marginal likelihoods requires the use of proper priors. In setting these priors it is advantageous to use priors that reflect our prior knowledge about the design of the experiment. In CEs we seldom have good prior expectations about the parameters of the logit, since the utility function is only identified up to a multiple of a constant. However, the logit model can be reparameterized so that the coefficients of the attributes reflect WTP for these attributes. We achieve this by estimating the model in WTP Space. This approach to model estimation has recently been adopted by several researchers within the literature eg, Train and Weeks (2005), Scarpa et al. (2008) and Balcombe et al. (2009).

If we assume that the first attribute is price (or payment) in the vector of attributes that define utility ($u_{ij} = x'_{ij}\beta + e_{ij}$) such that $x'_{ij} = \left(-p_{ij}, z'_{ij}\right)$ where p_{ij} is the payment and z'_{ij} are the other attributes, then the parameter vector can be expressed as:

$$\beta = \beta_0 \left(1, \gamma\right) \text{ where } \gamma = \left(\frac{\beta_1}{\beta_0}, \frac{\beta_2}{\beta_0} .... \frac{\beta_k}{\beta_0}\right), \tag{26}$$

where γ contains the estimates of WTP for each of the attributes. Axiomatically, we would also expect the value of β_0 to be positive. Therefore, a transformation of

$$\beta_0 = \exp\left(\gamma_0\right), \tag{27}$$

where γ_0 is unbounded, ensures that β_0 is positive. This reparameterization has no impact on the likelihood function under the assumption that the maximal value of the likelihood function is in the region where $\beta_0 > 0$. However, placing informative priors on the parameters γ_0 and γ may have quite a different effect than if priors are placed on γ. More importantly, we have some reasonable prior information as to the likely values of γ as these will be estimates of WTP. This discussed in detail below.

3 Choice Experiment Design and Model Estimation

3.1 Choice Experiment Design and Survey Returns

The CE employed in this paper was designed and implemented to examine consumers' WTP for reductions in the nutrients represented in the TLS (ie, Sugar, Fat, Saturates and Salt). The CE indicated the aggregate TLS for a hypothetical basket of goods as opposed to specific food items. We understood from the outset that this was a critical design issue. The initial piloting of the CE indicated that respondents could readily conceptualize a representative basket in terms of the aggregate TLS. So that the basket could be understood in terms of potential products we indicated to survey participants the typical sorts of goods that might be included in the basket. This choice of goods was based on research by Synovate (2005) for the FSA. Thus, we did not construct specific baskets of goods as we wished to avoid confusing items of food choice in the basket with the TLS.

We took this approach because although the TLS can and should be applied at the single product level it does need to be seen as a means by which to achieve a healthy diet for a mix of all food being consumed. There is nothing intrinsically wrong for a consumer to eat a bag of crisps, or a piece of chocolate, both of which will have Red labels for several nutrients, as long as they compensate for these food choices with moderation elsewhere in their diet. However, it is also clear from research that consumers find it hard to use nutritional label information to position specific food items within their overall diet (Cowburn and Stockley, 2005). Also Wansink and Chandon (2006) observe that consumers who select the health option then over compensate with less healthy options such that the net effect is a negative impact in their dietary intake. Thus, designing a CE that only considers a single food item could yield behavioral outcomes that do not capture how individuals should be employing the TLS as a means to achieve a healthy diet.

The final design of the CE was a choice card that presented the TLS nutrients using the associated colours as the main attributes along with the price of the basket. This meant we had four nutrients: Salt, Sugar, Fat and Saturates. Based on the mix of goods in the hypothetical basket we could deduce an appropriate Status Quo option to include in all the choice sets. The Status Quo option was defined as Amber for Salt, Sugar and Saturates and Red for Fat. We then established a price for the Status Quo basket of goods by referring to the National Statistics (2007) publication, Family Food in 2005-06. This produced a value of £20. We then determined the appropriate number of price points to include in the final design which was £15, £18, £25 and £30. We included prices below the Status Quo price because we considered it likely that some consumers are more price sensitive than health sensitive.

Given our set of attributes and the number of levels a full factorial design, ensuring a balance across attributes yielded 24 choice sets. Each of the 24 choice cards included the Status Quo option, and two other food baskets and the DK. We generated the food basket options by randomly pairing from the original set of 24. In the final design we decided not to label the options with the choice sets as we did not wish to explicitly signal to respondents the types of basket being considered. Finally, to ensure that we did not overload participants and to avoid response fatigue we divided the 24 choice sets into four groups of six with each respondent answering six choice sets. An example of a choice card is shown in Figure 1.

{Approximate Position of Figure 1}

The survey instrument was sent out in the mail to a random sample of 3,000 UK households in late 2007. It was a single shot survey. To induce participation all respondents who submitted a completed version of the survey where entered into a prize draw to win one of four gift vouchers worth £25. Overall we received 477 useable returns.

Overall, the average age of respondents is 48 years, compared to the UK average of 39 in 2007. The average age of respondents is only slightly higher than related consumer survey work on food labels by Loureiro et al. (2006) and Berning et al. (2008). Sample average income is £24,500 which compares to £30,000 for all households including retirees in 2006/07. Our returns were 81 percent females and 19 percent males which is an over-representation of females albeit only marginally higher than that reported in food related research. Some 73 percent of respondents are married and 67 percent of respondents have no children living at home. This compares to UK statistics which indicate that there are 33 percent single households and 60 percent of households with no children. Overall our sample had 0.6 children in the household. Finally, 60 percent of our respondents are employed and 30 percent 30 percent have achieved some level of university education.

3.2 The Model and Priors

The precise form of the model we estimate is as follows

$$
\begin{aligned}
(28) \qquad U &= \exp(\gamma_0) \times (-p + \gamma_1 d_{salt,g} + \gamma_2 d_{salt,r} + \gamma_3 d_{sug,g} \\
&\quad + \gamma_4 d_{sug,r} + \gamma_5 d_{fat,g} + \gamma_6 d_{fat,r} + \gamma_7 d_{sat,g} + \gamma_8 d_{sat,r}) + e
\end{aligned}
$$

where $p$is price, $d_{salt,g}$ and $d_{salt,r}$ are dummies which takes the value one if the salt nutrient attribute is Green and Red respectively, and the other dummies are similarly defined for the attributes sugar (sug) fat and saturate fat (sat). Since there are three colors, Red, Amber and Green, we would expect that each of the Red dummies would have negative coefficients as they reflect the WTP for moving from Amber to Red, and the Green dummies would have positive coefficients as they represent the WTP for moving from Amber to Green.

Given our experimental design and our model specification in (28) it follows that the parameters $\{\gamma_i : \ i = 1, ...8\}$ represent the WTP for a movement from Amber (the middle level) to Green (the healthier level) and Amber to Red (the less healthy option) for each nutrient. Our prior expectations are that a £5 difference in WTP (on a given nutrient) for Amber and other colors (Red or Green) would be a relatively large value. For example, this would imply that faced with the representative basket, or one where all nutrients were Red, survey respondents would (on average) only select this basket if it were less than £5. On the other hand it would imply that a move to a basket of all Greens relative to the status quo basket would be worth over £25. Consequently, we set the standard deviation on each of the WTP parameters $\{\gamma_i : \ i = 1, ...8\}$ equal to five (with a mean of zero). Given our expectations, this is a relatively diffuse prior, since it gives a considerable amount of mass above the £5 level. Since it is plausible that some nutrients may have higher WTPs than £5, we would not wish to use a value substantially less than five. Therefore, the WTP results presented in the empirical section of the paper use these priors.

Preliminary data analysis we found that the WTP estimates become larger if we use more diffuse priors. However, of more importance in the context of this research are the parameters that define respondents tendency to report DK (α, as defined in subsection on measuring utility). Therefore, we experimented with alternative priors, and the findings on the parameters describing the respondents tendency to report DK (α) are not changed in any substantive sense if we vary the priors for $\{\gamma_i\}$.

For the parameter γ_0 (in 28) we also use a standard deviation of five, but this translates into a highly diffuse prior given that $\beta_0 = \exp(\gamma_0)$. We note that the use of the normal prior for γ_0 implies that the prior for β_0 is log normal (an assumption that has been commonly used in WTP studies).

The appropriate priors on α_0 and α_1 are a little less easily deduced. However, a standard deviation for α_0 of five or more can be shown to be highly diffuse where $\alpha_1 = 0$, in the sense that it results in an approximately equal prior probability of 0.5, that $\theta_i = 1$ and $\theta_i = 0$ (as defined in equations 18, 19, and 20). Thus, with if we employ a more diffuse prior than this it has approximately the same implied prior precision for θ_i.

For the model we have developed it can easily be shown that the entropy measure is defined over the interval $(-\ln(3), 0)$. This emerges from the Shannon measure of entropy as we have $k = 3$. In this case it can also be shown that a standard deviation of five or more for α_1 will result in an approximately equal prior probability (close to 0.5) such that $\theta_i = 1$ and $\theta_i = 0$ regardless of the probability values of $\{p_{ij}\}$ or $\{p_{ij}^*\}$ (in equations 18, 19, and 20). This also follows for the jointly independent priors for α_0 and α_1 with standard deviations of five. Thus, as our most diffuse prior, we adopt standard standard deviations of five, but we experiment with smaller variances to show that the results are robust to these assumptions.

3.3 Estimation

Using the posterior densities, a random walk Metropolis-Hastings algorithm can be employed to map the posterior. When estimating our models we used a burn-in of 20,000 iterations followed by 500,000 further iterations in which each 10th draw was retained, and used for estimation of the mean and standard deviation of the posterior distributions of the marginal likelihoods. The marginal likelihoods were calculated using the method of Gelfand and Dey (1994). Convergence was monitored visually, along with modified t-tests for the difference of the first half and second half of the chain. All models appeared to converge well and accurately using the above settings.

4 Empirical Section

4.1 Results

In this section we begin by report the results for the three models previously defined. In Table 1 the logged marginal likelihoods of the three models estimated are presented. The results we present are for the priors on the WTP discussed in Section 3.

{Approximate Position of Table 1}

The first thing to note in Table 1 is that while the marginal likelihoods for the models are sensitive to the priors on the γ parameters, the Bayes Ratios for the three models (1,2,3) which have alternative treatments of the DKs, are quite insensitive to variations in the γ parameters. However, as discussed in the section on the choice of priors, variations on the priors for α may differ depending on the prior standard deviation. Accordingly, we first take a relatively diffuse prior (standard deviation 5) and then investigate whether a much tighter prior (standard deviation $\sqrt{5}$) has a significant effect on the marginal likelihoods (and Bayes Ratios). The results for the more diffuse prior are presented in the first column, followed by less diffuse prior in the second column. The Bayes Ratios that we present are relative to Model 1. Therefore, in the first row the Bayes Ratio is 1. The second and third Bayes Ratios reflect the prior odds that the Model 2 or 3 is preferred to Model 1 given equal prior odds. As can be seen from Table 1, the least preferred model (the one with the smallest logged marginal likelihood) is where the probability of replying DK is a constant that does not depend on the entropy measure (Model 1).

Of the two models that use the entropy measure the preferred model is the one where the entropy suggests that where the respondent is relatively indifferent to the top two options (Model 3). Models 2 and 3 have larger logged marginal likelihoods than Model 1, regardless of whether the relatively diffuse prior or less diffuse prior is adopted. For example, using the prior standard deviation of five, Model 2, has a posterior odds of 2.7:1 over Model 1. Model 3 is also preferred over Model 1 and Model 2. Therefore, these results indicate that DKs are indeed more common in sets that have options with similar levels of utility. Moreover, it is in circumstances when respondents are presented with options where the 2 most preferred are similar (and a third that they may not prefer), is most likely to result in a DK.

The coefficient estimates for the preferred model (Model 3, under the more diffuse priors) are presented in Table 2.

{Approximate Position of Table 2}

The results in Table 2 are the estimated mean and standard deviations of the posterior distributions. Median estimates were approximately the same as for the mean. The sign of each of the WTPs are in accordance with prior expectations. We can also observe that the rank order of the estimates indicates that respondents are WTP most to avoid Salt, Saturates, Fat and Sugar. Furthermore, from Table 2 we can see that the WTP for the movement from Amber to Red is larger (in absolute terms) than for Amber to Green for each of the four nutrient groups. This is to be expected as there is some anecdotal evidence to suggest that shoppers are showing a strong aversion to food products that have Red labels.

Turning to the issue of similaritywe can consider the parameter of most interest given the models developed in this paper, α_1. The negative value of -2.51 reported in Table 2 indicates that there is an increased tendency to report DK as the similarity of the responses increases As can be seen from Table 2, the posterior mass below zero for this parameters would be nearly unity. The distribution for this parameter is shown in Figure 2.

{Approximate Position of Figure 2}

From Figure 2 it is evident that the posterior mass lies in the negative region and is a symmetrical bell shaped distribution.

Next we consider the magnitude of the WTP estimates. These are large. For example, $\gamma_1 = 9.81$ suggests that, on average, consumers are WTP an additional £9.81 for the basket of goods if it has a Green rather than an Amber for Salt and £22.75 if it has Green rather than Red. It is also evident that respondents were able to differentiate between nutrient groups, with very different (non-overlapping) estimated WTP distributions. This result is especially strong when considering Reds. For example, for Salt, the WTP of £22.75 and a standard deviation 1.62 would give an 95% interval that would lie within (£19, £26) while for Sugar the 95% interval would lie within (£10, £16).

It is briefly worth considering why our WTP estimates are on the large side. As we explained the CE presented the survey respondents with a basket of shopping that was represented by the aggregate levels of the various nutrients plus price. We designed the CE this way to avoid undue complexity. However, it is possible that some of our respondents have interpreted a higher price as implying something about excluded attributes. For example, price could have been interpreted as a measure of quality. This form of behaviour has been identified in previous research by Brucks et al. (2000) especially for goods that exhibit credence attributes. A credence attribute is a product attribute that cannot be determined by the consumer during the search for a product or during consumption. The imprtance of credence attributes with respect to food purchase has been examined by a number of researchers (eg, Loureiro and Umberger, 2007). The type of attribute that have been considered include country-of-origin labeling, geograhical indications and various food safety systems such as traceability. If this is the case it will be necessary to control for this type of effect in any future research that examines the TLS.

5 Conclusions

This paper has introduced a new method of treating DK options within CE which have been included to capture respondent uncertainty. The need to develop CE that capture respondent uncertainty is only just beginning and in common with the CV literature there is a need to ensure that the data collected in this type of CE is used efficiently. Thus, the analysis presented in this paper provides the first attempt to reconcile the use of an option in a choice set to capture respondent uncertainty, an aspect of survey design and estimation that has generated many developments in the CV literature.

The econometric model developed in this paper works by conditioning the probability of a DK response on a similarity measure, in this case a Shannon entropy measure. By using the DK responses in this manner we have demonstrated there is the potential to utilize information contained in the DK responses rather than throwing potentially useful information away. This method has been applied to CE data that collected to examine consumer responsiveness to nutrient labelling on food items. Specifically, the TLS format that has been recommended by the UK Food Standards Agency.

Using Bayesian estimation and testing methods we found that DK responses were dependent on the similarity of the choices within the choice sets. This result implies that DK responses can be meaningfully incorporated into the estimation process. In particular, we found that when the two most preferred options were similar, this increased the chances that the respondent would reply DK. Our CE also revealed very high WTP estimates for changes in all nutrients, even when we employed informative priors that limited the propensity of the WTP estimates to be large. Our contention is that there is a strong normative component in the framing of choice sets using the TLS since in the minds of respondents there is an establish association with these colors: Red means stop and Green means go. Potentially, however, we believe that this effect may not only manifest itself within the experimental setting, but within real market transactions

Finally, the framework for treating DKs used in this study could in principle be extended to the mixed logit, which is an increasingly popular way to estimate the parameters of choice models. However, our hypothesis that similarity in choice sets could also be more extensively explored by the deliberate design of choice sets that were clearly differentiated according to whether the options within the choice sets were similar or dissimilar in terms of the utility that they provided. It should also be noted that there is an indirect relationship between the methods developed in this paper and those that are employed in the design of choice sets generally. There is a growing literature that is concerned with the efficient design of choice sets (see Scarpa and Rose, 2008) so as to reduce the necessary sample size of a survey whilst achieving the desired level of estimation accuracy. Within this literature design efficiency frequently ignores behavioral efficiency which relates to the way in which survey respondents actually relate to the CE in practice. However, if survey design methods generate choice sets that yield similar options we would contend that this is likely to lead to increased uncertainty on the part of survey respondents. Thus, in practice efficiency of survey design and respondent behavior should be considered simultaneously.

References

Balcombe, K.G., A. Chalak, and I.M. Fraser (2009). Model Selection for the Mixed Logit with Bayesian Estimation, *Journal of Environmental Economics and Management,* 57(2): 226-237

Balcombe, K.G. and I.M. Fraser (2009). Dichotomous Choice Contingent Valuation with Don't Know Responses and Misreporting. *Journal of Applied Econometrics,* Forthcoming.

Banzhaf, M.R., Johnson, F.R. and K.E. Mathews. (2001). *Opt-Out Alternatives and Anglers' Stated Preferences*, Chapter 8, pages 157-177, in Bennett, J. and Blamey, R. (Edited), *The Choice Modelling Approach to Environmental Valuation*, Edward Elgar.

Berning, J.P., Chouinard, H. and J.J. McCluskey. (2008). Consumer Preferences for Detailed Versus Summary Formats of Nutrition Information on Grocery Store Shelf Labels. *Journal of Agricultural and Food Industrial Organization* Volume 6, (Article 6), www.bepress.com/jafio/vol6/iss1/art6.

Brucks, M., V.A. Zeithaml and G. Naylor (2000). Price and Brand Name as Indicators of Quality Dimensions for Consumer Durables, *Journal of the Academy of Mrketing Science*, 29(3): 359-374.

Cowburn, G. and Stockley, L. (2005). Consumer Understanding and Use of Nutrition Labelling: A Systematic Review, *Public Health Nutrition*, 8(1): 21-28.

Dellaert, B.G.C., Brazell, J.D. and J.L. Louviere (1999). The Effect of Attribute Variation on Consumer Choice Consistency. *Marketing Letters* 10(2): 139-147.

DeShazo, J.R. and G. Fermo. (2002). Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management* 44: 123-143.

Evans, M.F. Flores N.E. and K.J. Boyle. (2003). Multiple-Bounded Uncertainty Choice Data as Probabilistic Intentions. *Land Economics* 79 (4): 549-560.

Feinchel, E.P., Lupi, F., Hoehn, J.P. and M.D. Kaplowitz. (2009). Split-Sample Tests of "No Opinion" Responses in an Attribute Based Choice Model. *Land Economics* 85(2): 349-363.

Gao Z, and T.C. Schroeder. (2009). Effects of Label Information on Consumer Willingness to Pay for Food Attributes. *American Journal of Agricultural Economics* 91(3) 795-809.

Gelfand, A. and D. Dey. (1994). Bayesian Model Choice: Asymptotics and Exact Calculations. *Journal of the Royal Statistical Society Series* B, 56: 501-504.

Golan, A., Judge, G. and D. Miller (1996). *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, John Wiley and Sons, Chichester, UK.

Haab, T.C. and K.E. McConnell. (2002). *Valuing Environmental and Natural Resources. The Econometrics of Non-Market Valuation* Edward Elgar, Cheltenham, UK.

Hatton MacDonald, D., M. Barnes, J. Bennett, M. Morrison and M.D. Young (2005). Using a Choice Experiment Modelling Approach for Customer Service Standards in Urban Water, *Journal of the American Water Resources Association*, 41(3): 719-728.

Hensher, D. A., J.M.Rose and W.H. Greene (2005). *Applied Choice Analysis. A Primer*, Cambridge University Press, Cambridge, UK.

Huber, J. and K. Zwerina. (1996). The Importance of Utility Balance in Efficient Choice Designs, *Journal of Marketing Research*, 33(3): 307-317..

Kontoleon, A. and M. Yabe. (2003). Assessing the Impacts of Alternative "Opt-Out" Formats in Choice Experiment Studies: Consumer Preference for Genetically Modified Content and Product Information in Food. *Journal of Agricultural Policy Research* 5: 1-43.

Kosenius, A-K. (2009).Causes of Response Uncertainty and Its Implications for WTP Estimation in Choice Experiments, Discussion Paper No. 29, Department of Economics and Management, University of Helsinki, Finland.

Loureiro, M., Gracia, A. and Nayga, R. (2006). Do Consumers Value Nutritional Labels, *European Review of Agricultural Economics*, 33(2): 249-268.

Loureiro, M.L. and W.J. Umberger (2007). A Choice Experiment Model for Beef: What US Consumer Responses Tell Us About Relative Preferences for Food Safety, Country-of0Origin Labeling and Traceability, *Food Policy*, 32: 496-514.

Lundhede, T.H., S.B. Olsen, J.B. Jacobsen and B.J. Thorsen (2009). Handling Respondent Uncertainty in Choice Experiments: Evaluating Recoding Approaches Against Explicit Modelling of Uncertainty, *Journal of Choice Modelling*, Forthcoming.

National Statistics. (2007). Family Food in 2005-06, A National Statistics Publication by DEFRA, London.

Scarpa, R. and J.M. Rose. (2008). Design Efficiency for Non-Market

Valuation with Choice Modelling: How to Measure it, what to Report and Why. *Australian Journal of Agricultural and Resource Economics* 52(3): 253-282.

Scarpa, R., Thiene, M. and Train, K. (2008). Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps, *American Journal of Agricultural Economics*, 90(4):994-1010

Synovate. (2005). Quantitative Evaluation of Alternative Food Signposting Concepts, Reported prepared for the FSA, November.

Train, K. and M. Weeks (2005). Discrete Choice Models in Preference Space and Willingness-to-Pay Space, in *Applications of Simulation Methods in Environmental Resource Economics*, R. Scarpa and A. Alberini (Eds.), Kluwer Academic Publishing.

Wansink, B. and Chandon, P. (2006). Can "Low-Fat" Nutrition Labels Lead to Obesity?, *Journal of Marketing Research*, XLIII: 605-617.

Welsh, M.P. and G.L. Poe. (1998). Elicitation Effects in Contingent Valuation: Comparisons to a Multiple Bounded Discrete Choice Approach. *Journal of Environmental Economics and Management* 36:170-185.

Table 1.Marginal Likelihoods

	Log Marginal Likelihoods		Bayes Ratios (v Model 1)	
	Prior stdv*=5	Prior stdv=$\sqrt{5}$	Prior stdv=5	Prior stdv=$\sqrt{5}$
Model 1	-2766.58	-2765.79	1	1
Model 2	-2765.59	-2764.12	2.7	5.3
Model 3	-2763.55	-2762.44	20.75	28.3

*The Prior Standard Deviations are for the parameters α_0 and α_1

Table 2: Coefficient Estimates (WTPs) for Model 3

	Mean	Standard Deviation
γ_0	-0.0317	0.0058
$\gamma_1(d_{salt,g})$	9.81	0.94
$\gamma_2(d_{salt,r})$	-22.75	1.62
$\gamma_3(d_{sug,g})$	10.30	1.13
$\gamma_4(d_{sug,r})$	-13.04	1.07
$\gamma_5(d_{fat,g})$	10.83	1.27
$\gamma_6(d_{fat,r})$	-14.17	1.27
$\gamma_7(d_{sat,g})$	12.49	1.03
$\gamma_8(d_{sat,r})$	-19.56	1.45
α_0	-4.88	0.47
α_1	-2.51	0.81

CHOICE CARD 1				
Food Basket	***Option 1***	***Option2***	***Option3***	**Don't Know**
Salt	Amber	Red	Green	
Sugar	Amber	Green	Amber	
Fat	Red	Amber	Red	
Saturates	Amber	Amber	Red	
Price of basket	£20	£25	£30	
Tick ONE and only one box				

Figure 1: Example Choice Card

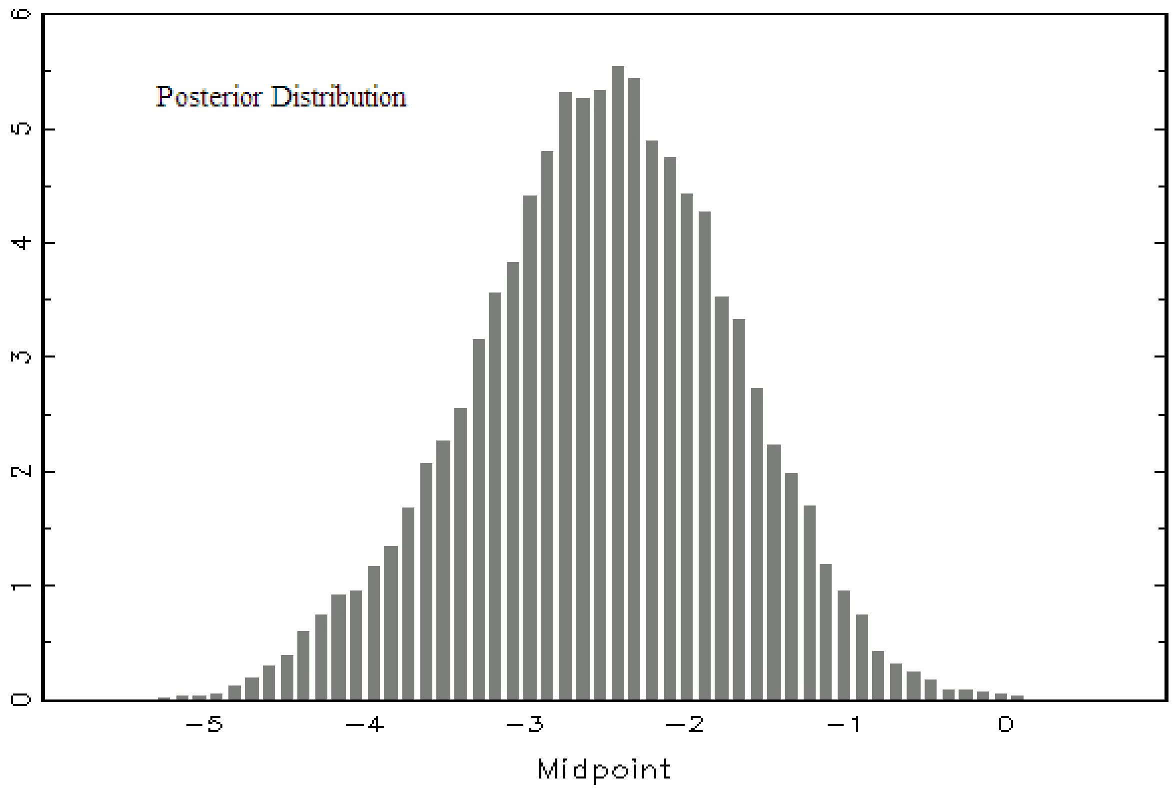

Figure 2:Posterior Distribution for α_1