The 'no-choice' alternative in conjoint choice experiments

Rinus Haaijer,* Wagner Kamakura† and Michel Wedel*
*University of Groningen and †University of Iowa

Conjoint choice designs are frequently applied in practice, and often a base alternative is added to the design. When such a 'no-choice' base alternative is present in conjoint choice experiments a constant term should be added to the design ('X'-) matrix with attribute dummies when effects type and/or linear coding is used for the attribute levels. Not including such a constant may result in a much lower model and predictive fit and even biased estimates for the (linear) attributes.

Introduction

Choice experiments have become prevalent as a mode of data collection in conjoint analysis in applied research. Their popularity seems due to their ability to mimic real market decisions. Experimental choice analysis has a number of advantages over traditional conjoint analysis, including that there are no differences in scale usage across respondents, that the choice tasks have higher external validity being more akin to the behaviour respondents display in reality, and that assumptions made in designing choice simulators can be avoided. The availability of new computer-assisted data-collection methods, in particular CBC from Sawtooth Software, has also greatly contributed to its popularity. In conjoint choice experiments respondents make choices from several sets of alternatives. Respondents choose one profile from each of several choice sets. In order to make the choice more realistic, in many conjoint experiments one of the alternatives in the choice sets is a 'no-choice' or 'none' option. This option can entail a real no-choice alternative ('None of the above') or an 'own-choice' alternative ('I keep my own product'). This base alternative, however, presents the problems of how to include it in the design of the choice experiment and how to accommodate it in the choice model. Regular choice alternatives are most often coded in the data matrix with effects-type or dummy coding. Since the no-choice alternative does not possess any of the attributes in the design, one may be tempted to code it simply as a series of zeros, which makes the fixed part of its utility zero in each choice set. In this paper we investigate several models that can be used to accommodate the no-choice option and we show that it is best to add a constant to the design or X-matrix of the conjoint experiment for the no-choice alternative when effects-type coding and/or linear coding is used for all attributes. Not including this constant leads to reduced fit and biased estimates of linear attributes.

The base alternative

In conjoint choice experiments a base alternative is included in the design of the experiment, among others, to scale the utilities between the various choice sets. A base alternative can be specified in several ways. First, it can be a regular profile that is held constant over all choice sets. Second, it can be specified as 'your current brand', and third, as a 'none', 'other' or 'no-choice' alternative (e.g. Louviere & Woodworth 1983; Batsell & Louviere 1991; Carson *et al.* 1994). Additional advantages of including a 'no-choice' or 'own' base alternative that are mentioned in the literature are that it would make the choice decision more realistic and would lead to better predictions of market penetrations. A disadvantage of a no-choice alternative is that it may lead respondents to avoid difficult choices, which detracts from the validity using the no-choice probability to estimate market shares. However, Johnson & Orme (1996) claim that this seems not to happen in conjoint choice experiments. In addition, the no-choice alternative gives no information about preferences for attributes of the choice alternatives, which is the main reason for doing a conjoint choice experiment. Dhar (1997) gives an overview of why and when respondents may choose a no-choice option in general. He states that respondents may choose the no-choice when none of the alternatives appears to be attractive, or when the decision-maker expects to find better alternatives by continuing to search. Furthermore, when subjects are uncertain about the range of potential alternatives they may continue to look for better alternatives and choose the 'none' in the early stages of a choice process. Dhar (1997) also shows that adding an attractive alternative to an already attractive choice set increases the preference of the no-choice option and adding an unattractive alternative to that choice set decreases the preference of the no-choice. This implies that when alternatives are close to each other in preference, people will choose the no-choice more often than when there is a clearly dominant or unattractive profile in the choice set. Huber & Zwerina (1996) stated that conjoint choice sets that are utility balanced are more informative and efficient compared to sets with dominant alternatives. However, according to the above, adding a no-choice alternative to such a choice set may influence this negatively, because people may avoid making the difficult choice. This is related to findings in the psychological literature, where it has been found that people prefer consequences that arise from inaction over those arising from action since the decision to stay within a status quo has certain psychological advantages (Baron & Ritov 1994; Dhar 1997). This may make people choose 'none' if there is no dominant alternative present in the choice set. The above shows that the reasons to choose the no-choice may be different from choosing any of the other 'real' profiles in a conjoint choice experiment. In this view, 'no-choice' cannot be seen as just another choice alternative, leading to potential violations of IIA, among others.

In this paper we investigate the no-choice option from a modelling point of view.[1] We begin by discussing a number of alternative model formulations. First, simply having a series of zeros describing the attribute values of the no-choice alternative seems a straightforward option, but this formulation may produce misleading results. When no-choice is coded as a series of zeros, its fixed part of utility is equal to zero by definition. However, when there are linear attributes present, these zero values of the no-choice alternative act as real levels of the linear attributes. For instance, when price is a linear attribute in the design, the zero value for no-choice will correspond to a zero price. We hypothesise that this can lead to a biased estimate of the parameter of the linear attribute.

Second, when all attributes are modelled with effects-type coding the bias discussed above does not arise, because all part-worths are now specified relative to the zero utility of the no-choice alternative. However, even when all attributes are coded with effects-type dummies, adding such a constant may improve model fit. This can be explained because the no-choice option in fact adds one level to all attributes. If there are, for example, S attributes each with L levels, then there are L^S possible alternatives and L part-worths to be estimates for each of the S attributes in the analysis. The inclusion of a no-choice option would produce a

[1] In the remainder of the paper we only mention the 'no-choice' (or 'none'), but the results also apply to the 'own' alternative when nothing is known about its characteristics to the researcher. Even when the (design) attributes and levels of the 'own' alternatives are known, there may be additional attributes, which have some utility to the consumer but which are not in the design of the experiment. In this case an extra constant in the design may act as some 'threshold' value for the 'own' alternative.

design with S attributes each with $L + 1$ levels. However, the resulting full $(L + 1)^S$ set of profiles cannot be constructed, since the $(L + 1)$th level of all attributes is confounded with the no-choice alternative. Thus, instead of $(L + 1)^S$ alternatives only $L^S + 1$ alternatives are possible. This is accommodated in the Multinomial Logit analysis by estimating the L part-worths (through $L - 1$ effects-type dummies) for each of the S attributes plus one part-worth for the no-choice option. Thus adding the no-choice constant is hypothesised to increase model fit. This change in the design of the choice experiment is hypothesised to reduce the potential bias in the estimates for linear attributes. Although this additional constant increases the number of parameters by one, it sets the utility level of the no-choice alternative, as explained above, and therefore compensates for the bias in the estimates for linear attributes in the design.[2]

Finally, another way to model the presence of a no-choice option is by specifying a Nested Logit model. When two nests are specified, one containing the no-choice and the other the real profiles, the no-choice alternative is no longer treated as just another alternative. The idea is that respondents first decide whether or not to choose and only when they decide to choose a real profile do they select one of them, leading to a nested choice decision. This way of modelling the no-choice potentially also removes the effects of linear attributes because the zeros of the no-choice are no longer treated as real levels, because they are now captured in a different nest. The Nested Logit model can also be tested for the situation with or without the presence of linear attributes next to effects-type coded attributes.

No-choice models

We use three models: (1) the Multinomial Logit model (MNL), (2) the Nested Multinomial Logit model (NMNL), and (3) the No-choice Multinomial Logit model. The difference between the Logit model and the No-choice Logit model is the extra constant (c_{nc}) added for the no-choice option in the design, but both models fall within the standard multinomial Logit context for conjoint experiments. In the Nested Logit model there is one extra parameter (λ) called the *dissimilarity coefficient* (Börsch-Supan 1990). When its value is equal to 1, the Logit and Nested Logit model are

[2] Note that the same can be accomplished by coding one of the attributes (e.g. brand) with dummy coding instead of effects-type coding. This also increases the number of parameters with one and sets the level of the no-choice. Another possibility is to add a constant for the 'real' alternatives instead and keep a series of zeros for the no-choice; in that case the constant is equal to $-c_{nc}$.

equal. In the Nested Logit model we assume that there are two nests, one containing the no-choice alternative, the other containing the remaining real alternatives.

In a conjoint choice model each respondent has to choose one alternative from each of several choice sets. These choice sets are constructed by dividing the total set of profiles over K choice sets. We assume that each choice set contains the same number of alternatives without losing generality. The utility of alternative m in choice set k for individual j is defined as:

$$U_{jkm} = X_{km}\beta + e_{jkm} \tag{1}$$

where X_{km} is a $(1 \times S)$ vector of variables representing characteristics of the mth choice alternative in choice set k. In most conjoint choice experiments no individual characteristics are present, so X does not depend on j. Note, however, that when an individualised design is used, X does depend on j, but we omit this index here for convenience. β is a $(S \times 1)$ vector of unknown parameters, and e_{jkm} is the error term. The MNL model treats observations coming from the same respondent as independent observations, and falls within the standard random utility approach. The choice probabilities p_{km} that alternative m is chosen from set k in the conjoint MNL approach can be obtained as (see for example Ben-Akiva & Lerman 1985):

$$p_{km} = \frac{\exp(X_{km}\beta)}{\sum_{n=1}^{M}\exp(X_{kn}\beta)} \tag{2}$$

The log-likelihood in the conjoint context is extended by adding a sum over choice sets to the standard random utility log-likelihood and is equal to:

$$l = \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{m=1}^{M} y_{jkm} \ln (p_{km}) \tag{3}$$

The probabilities in a Nested Logit model are different from the probabilities in the standard MNL model. Assume that there are N nests and each nest n contains M_n alternatives (the index representing the choice sets is suppressed for the moment), then the probability that alternative m

is chosen from nest n is calculated as (cf. e.g. McFadden 1981):

$$P(n, m) = P(m|n) \bullet P(n) \tag{4}$$

where

$$P(m|n) = \exp(X_{nm}\beta) / \sum_{m'=1}^{M_n} \exp(X_{n'm}\beta) \tag{5}$$

is the standard MNL probability *within* a nest, and

$$P(n) = \exp(\lambda V_n) / \sum_{n'=1}^{N} \exp(\lambda V_{n'}) \tag{6}$$

is the probability of choosing nest n, λ is the above-mentioned dissimilarity coefficient, and $V_n = \ln[\exp(\sum_{m=1}^{M_n} X_{nm}\beta)]$ is the *inclusive value* for nest n. When $\lambda = 1$ it is easy to show that the Nested Logit is equal to the Logit model. When there is only one nest ($N = 1$) (5) reduces to the standard Logit formula (2) for conjoint experiments. In the case of the Nested Logit model, when the no-choice is coded as a series of zeros for the attributes, the inclusive value (V_n) of the nest containing only the no-choice is equal to zero in the Nested Logit model. Furthermore, $P(m|n) = 1$ for the nest containing the no-choice because it contains only one alternative.

Empirical investigations of modelling options

The following provides two applications to commercial conjoint choice data to illustrate the relative fits of the alternative models and coding of the attributes.

First application

We describe the results of the models outlined in the previous section, using a conjoint choice experiment on a technological product with six attributes:[3] brand (six levels), speed (four levels), technology type (six levels), digitising option (no and two yes levels), facsimile capable (y/n), and price (four levels). For all models we use three versions: in the first situation the price and speed attributes are coded linear with {1, 2, 3, 4}

[3] We thank Rich Johnson from Sawtooth Software for allowing us to analyse these data. Because the data are confidential we cannot give more details about the attributes and levels.

for the four levels respectively (speed ascending, price descending) and the other attributes are coded using effects-type coding. In the second situation we use as linear codes {–3, –1, 1, 3} for the linear levels instead, to investigate whether mean centring the linear levels solves (part of) the problem. In the third situation all attributes are coded with effects-type coding. We used 200 respondents who each had to choose from 20 choice sets with four alternatives, where the last alternative is the 'no-choice' option, which is defined as none of the above alternatives. We used the first 12 choice sets for estimation and the last eight for prediction purposes. Each respondent had to choose from individualised choice sets. We compare the results of the models on the log-likelihood value, AIC (Akaike 1973) and BIC statistics (Schwarz 1978) and the Pseudo R^2 value (e.g. McFadden 1976) relative to a null-model in which all probabilities in a choice set are equal to $1/M$. The AIC criterion is defined as: $AIC = -2 \ln L + 2n$, where n is the total number of estimated parameters in the model and the BIC criterion is defined as: $BIC = -2 \ln L + n \ln (O)$, where O is the number of independent observations in the conjoint choice experiment. We test differences in the likelihood values for models that are nested with the likelihood ratio (LR) test.

In Table 1 the estimation results are listed for all models with linear attributes; the left-hand side gives the results for the models with linear levels {1, 2, 3, 4}, the right-hand side with linear levels {–3, –1, 1, 3} (the results for the models where all attributes are coded with effects-type dummies are not shown). Note that when effects-type coding is used for an attribute, the part-worth for the last level of that attribute can be obtained by taking the sum of the estimates of the other levels of that attribute and changing the sign. Note also that the Nested Logit and the No-choice Logit models are not nested, but both are nested within the standard Logit model. In the nested Logit model we do not estimate λ itself but estimate $(1 - \lambda)$ instead, to have a direct test on $\lambda = 1$.

The first conclusion that can be drawn from Table 1 is that the No-choice Logit model gives the best overall fit, in both situations, and converged to the same point. The likelihood is significantly better than the standard MNL model (LR(1 df) tests, $p < 0.01$). The No-choice Logit model and the Nested Logit model are not nested, so these models cannot be compared with an LR test. The *AIC* and *BIC* values show, however, that the No-choice Logit model fits better than the Nested Logit model, which itself is significantly better than the Logit model (LR(1) tests, $p < 0.01$) again in both situations. Table 1 also shows that the estimates for the dissimilarity coefficients (λ) are significantly different from 1 for

Table 1 Estimation and prediction results: first application

	Levels linear attributes {1, 2, 3, 4}						Levels linear attributes {−3,−1,1,3}					
Model	MNL		NMNL		No-choice MNL		MNL		NMNL		No-choice MNL	
Parameter	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e
Brand A β_{01}	0.386	0.061*	0.522	0.077*	0.472	0.066*	0.350	0.064*	0.524	0.076*	0.472	0.066*
Brand B β_{02}	−0.013	0.067	−0.009	0.078	−0.011	0.070	−0.012	0.069	−0.010	0.078	−0.011	0.070
Brand C β_{03}	0.052	0.067	−0.006	0.080	0.037	0.071	0.024	0.071	0.001	0.080	0.037	0.071
Brand D β_{04}	−0.150	0.071*	−0.166	0.087*	−0.187	0.075*	−0.142	0.074*	−0.176	0.086*	−0.187	0.075*
Brand E β_{05}	0.011	0.067	−0.018	0.081	0.020	0.071	0.013	0.070*	−0.011	0.080	0.021	0.071
Speed β_{06}	−0.237	0.021	0.118	0.031*	0.129	0.028*	0.046	0.014*	0.067	0.016*	0.064	0.014*
Tech. type A β_{07}	−0.531	0.081*	−0.678	0.096*	−0.633	0.086*	−0.451	0.084*	−0.682	0.095*	−0.634	0.086*
Tech. type B β_{08}	0.505	0.060*	0.613	0.074*	0.575	0.064*	0.432	0.062*	0.621	0.073*	0.575	0.064*
Tech. type C β_{09}	−0.321	0.075*	−0.366	0.086*	−0.368	0.078*	−0.281	0.077*	−0.377	0.086*	−0.368	0.078*
Tech. type D β_{10}	0.514	0.060*	0.635	0.074*	0.628	0.064*	0.471	0.062*	0.651	0.073*	0.628	0.064*
Tech. type E β_{11}	−0.132	0.071	−0.149	0.084	−0.173	0.075	−0.132	0.073	−0.159	0.083	−0.173	0.075
Dig. opt (n) β_{12}	−0.586	0.049*	−0.714	0.055*	−0.732	0.052*	−0.488	0.050*	−0.721	0.055*	−0.732	0.052*
Dig. opt (y1) β_{13}	0.128	0.041*	0.180	0.046*	0.172	0.043*	0.099	0.043*	0.178	0.046*	0.172	0.043*
Facsimile β_{14}	−0.445	0.030*	−0.528	0.035*	−0.543	0.032*	−0.386	0.031*	−0.539	0.035*	−0.543	0.032*
Price β_{15}	−0.013	0.020	0.385	0.031*	0.396	0.028*	0.144	0.014*	0.203	0.015*	0.198	0.014*
Nested Logit $1-\lambda$	–	–	0.924	0.009*	–	–	–	–	0.840	0.017*	–	–
No-choice c_{nc}	–	–	–	–	2.461	0.121*	–	–	–	–	1.150	0.048*
Fit statistics												
Ln-likelihood		−2906.738		−2715.413		−2663.017		−2947.716		−2701.927		−2663.017
AIC		5843.476		5462.826		5358.035		5925.431		5435.854		5358.035
BIC		5930.224		5555.358		5450.566		6012.180		5528.386		5450.567
Pseudo R^2		0.126		0.184		0.200		0.114		0.184		0.200
Prediction statistics												
Ln-likelihood		−1883.069		−1741.723		−1706.087		−1960.978		−1735.014		−1706.087
AIC		3796.139		3515.448		3444.175		3951.956		3502.028		3444.175
BIC		3876.805		3601.492		3530.219		4032.622		3588.072		3530.219
Pseudo R^2		0.151		0.215		0.231		0.116		0.218		0.231

$^*p < 0.05$

the Nested Logit model, hence the Nested Logit differs significantly from the MNL model.

When the β-estimates are compared, Table 1 shows that the parameters estimates of the attributes with a dummy coding ($\beta_1,\ldots,\beta_5$, $\beta_7,\ldots,\beta_{14}$) are somewhat different, although not dramatically so. However, on the left-hand side of Table 1, the coefficients of the linear attributes (β_6, β_{15}) in the standard MNL model differ strongly from the other two models. Whereas the estimate for speed is negative (a high level is unattractive) and significant for the MNL model, it is positive (a high level is attractive) and significant for the other models. The price parameter shows a similar effect; it is negative but not significant in one situation and positive in the other for the MNL model, and positive (lower price is more attractive) and significant in the other two models. Clearly, both estimates for the linear attributes show a strong negative bias. Note, however, that there are also differences in the other part-worth estimates across the models. The right-hand side of Table 1 shows that when the speed and price variables are coded with values such that the mean of the levels is zero (the same can of course be obtained by mean-centring the linear levels on the left-hand side of Table 1), the estimates for speed and price no longer show the wrong sign, but are still biased downwards compared to the other models. In the No-choice MNL model the estimates for all parameters are equal in both situations, except for the linear parameters which, on the right-hand side of Table 1, have exactly half the value of those on the left-hand side, which is the result of the doubled step-length of the linear levels. When the attributes price and speed are also coded with effects-type coding (not shown), the β-estimates are more similar across the three models, having the right signs.

The conclusion that can be drawn from the above results is that the presence of a no-choice alternative and linearly coded attributes can give very misleading results, in particular for the parameters of those linear attributes when the conjoint choice data are estimated with a standard Logit model. However, the parameters of attributes coded with effects-type dummies are also affected, be it less severely. When all attributes are coded with effects-type coding the bias seems less strong, but coefficients estimates are still highly attenuated. Overall fit can be improved substantially by specifying a Nested Logit or by adding a No-choice constant to the design. When we compare the Nested Logit and the No-choice Logit results we see that both compensate for the no-choice zero level for the linear attributes, but there are some differences in the magnitudes of the estimated coefficients – some in the range 5–10% –

which may be important in substantive interpretation. However, the fit of the No-choice Logit model is much better than the Nested Logit model. Note that in Table 1 the estimates for the no-choice constant are relatively large and positive. This means that the no-choice has a high overall utility, which is also shown by the number of times the no-choice alternative was actually chosen (in 43.1% of all choice sets).

The estimates in Table 1 were used to predict the eight holdout choice sets. Table 1 also gives the values of the statistics for the predictive fit of the three models for the three different design options considered. The No-choice Logit model gives the best predictions which are significantly better than the Logit model (LR(1) tests, $p < 0.01$) and which are also better than the Nested Logit model in all situations. The Nested Logit model also predicts significantly better than the standard Logit model (LR(1) tests, $p < 0.01$). Thus the predictive validity results confirm the results on model fit. Note that although the MNL model with linear levels {1, 2, 3, 4} is clearly misspecified (as could be seen from the speed and price estimates in Table 1), the likelihood, both in estimation and prediction, is better than those of the MNL models with the two other ways of coding. However, in all situations the Nested Logit and No-choice Logit models show superior fit. We consider two further extensions.

First, in conjoint studies first-order brand–price interaction effects are often included in the design. The question is whether the above results change when such interactions are included. We also tested what the effects on these interaction effect parameters are for the various models and design specifications. In almost all cases no significant interaction terms were obtained and no (significant) improvement of the likelihood was found.

Second, one could expect combination of the No-choice and Nested MNL models, having two more parameters than the standard MNL model to provide a better fit and predictive validity than the nested and no-choice dummy MNL model specifications. This model matches the data, but converges to the same log-likelihood as the No-choice MNL model and is not identified, as the Hessian matrix of second-order derivatives has a zero eigenvalue. Therefore, we conclude that the No-choice MNL model remains the best-fitting model.

Second application

We describe the results of a second application on a technological product with six attributes: brand (five levels), performance (three levels), way of ordering (three levels), warranty (three levels), service (three levels), and price (four levels). For all models we use two versions; in the first situation the price, warranty and performance attributes are coded linear with warranty and performance: {1, 2, 3} and price: {1.5, 1.8, 2.2, 2.5}, respectively, and the other attributes are coded using effects-type coding. In the second situation we use mean centring so that the linear codes are warranty and performance: {–1, 0, 1} and price: {–0.5, –0.2, 0.2, 0.5}. We used 200 respondents who each had to choose from eight choice sets with four alternatives, where the last alternative is the 'no-choice' option, which is defined as none of the above alternatives. Each respondent had to choose from individualised choice sets. In this case we have no holdout choice sets available. We compare the results of the models on the log-likelihood value, AIC and BIC statistics and the Pseudo R^2 value relative to a null-model in which all probabilities in a choice set are equal to $1/M$, as above.

The estimation results are listed for all models in Table 2. The first conclusion that can be drawn from Table 2 is that the No-choice Logit model again gives the best overall fit in both cases with different coding of the variables. The likelihood is significantly better than the standard MNL model in all cases (LR(1 df) tests, $p < 0.01$) and the *AIC* and *BIC* values show that it fits better than the Nested Logit model, which itself is as good as the Logit model in this application. It is surprising that for this dataset, if the linear attributes are not mean-centred, the nested MNL converges to the same solution as the MNL model. However, the no-choice logit does better on all statistics in both cases. The estimates for the no-choice constant are significant in both cases.

Table 2 shows that the coefficients of the linear attributes, price and performance, in the standard MNL are again affected by the mean-centring operation. The estimates for the linear attributes are biased. Note, however, that there are also differences in the other part-worth estimates across the models, albeit they are smaller. In the No-choice MNL model the estimates for all parameters are equal across the different situations: only the magnitude of the no-choice constant is affected by the different coding of the linear attributes. Again, this is a desirable feature of this model.

The conclusion is that the results for this application replicate those of the first application shown above. The presence of a no-choice alternative and linearly coded attributes can give misleading results, in particular for the parameters of those linear attributes when the conjoint choice data are estimated with a standard Logit model. Fit is substantially improved by adding a no-choice constant to the design matrix.

Table 2 Estimation and prediction results: second application

Linear levels:	Performance {1, 2, 3}, warranty {0.5, 1, 5}, price {1.5, 1.8, 2.2, 2.5}						Performance, warranty and price: mean-centred					
Model	MNL		Nested MNL		No-choice MNL		MNL		Nested MNL		No-choice MNL	
Parameter	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e	Est.	s.e
Brand A	0.349*	0.069	0.349*	0.069	0.342*	0.069	0.360*	0.069	0.360*	0.069	0.342*	0.069
Brand B	0.193*	0.072	0.193*	0.072	0.192*	0.072	0.203*	0.072	0.203*	0.072	0.192*	0.072
Brand C	0.149*	0.069	0.149*	0.069	0.147*	0.069	0.155*	0.069	0.155*	0.069	0.146*	0.069
Brand D	−0.165*	0.072	−0.165*	0.072	−0.162*	0.073	−0.169*	0.072	−0.169*	0.072	−0.162*	0.073
Performance	1.157*	0.045	1.157*	0.050	1.093*	0.049	1.170*	0.047	1.170*	0.050	1.093*	0.049
Order phone	−0.048	0.047	−0.048	0.047	−0.046	0.047	−0.051	0.047	−0.051	0.047	−0.046	0.047
Order store	−0.008	0.046	−0.008	0.046	−0.008	0.046	−0.008	0.046	−0.008	0.046	−0.008	0.046
Warranty	0.173*	0.016	0.173*	0.016	0.162*	0.016	0.173*	0.016	0.173*	0.016	0.162*	0.016
Serv. ship back	−0.639*	0.052	−0.639*	0.052	−0.626*	0.052	−0.673*	0.052	−0.673*	0.053	−0.626*	0.052
Serv. dealer	0.171*	0.046	0.171*	0.046	0.170*	0.047	0.184*	0.046	0.184*	0.046	0.170*	0.047
Price	−1.220*	0.067	−1.220*	0.084	−1.486*	0.097	−1.579*	0.096	−1.579*	0.098	−1.486*	0.097
Nested Logit $1 - \lambda$	–	–	0.000	na	–	–	–	–	0.000	na	–	–
No-choice	–	–	–	–	−0.807*	0.215	–	–	–	–	−0.358*	0.081
Statistics												
Ln-likelihood		−1569.002		−1569.002		−1561.606		−1571.630		−1571.630		−1561.606
AIC		3160.004		3162.004		3147.211		3165.260		3167.260		3147.211
BIC		3219.159		3226.537		3211.744		3224.416		3231.794		3211.744
Pseudo R^2		0.293		0.293		0.296		0.291		0.291		0.296

*Significant 5%

Conclusions and discussion

Respondents may choose the no-choice alternative for two reasons. First, they may not be interested in the product category under research and for this reason choose the no-choice. In such a situation they would first decide whether or not to choose the offered product profiles and the Nested MNL model may be the most appropriate specification to describe this behaviour, since it puts the no-choice alternative in a different nest from these product profiles. The probability of the no-choice alternative may be an indication of the overall preference of the product in this case, and the model may be used to obtain an estimate of the overall attractiveness of the product category. Thus in this case the no-choice alternative would capture 'real' behaviour of consumers in the marketplace. Second, respondents may choose the no-choice because no real alternative in the choice set is attractive enough or because all alternatives are equally attractive and they do not want to waste more time on making the difficult choice. In this case the respondent treats the no-choice as 'just another' alternative, and the no-choice captures an effect specific to the task. If this is the case, the MNL model, with a no-choice constant, is the appropriate model to use, since it treats all alternatives equally. Now the utility of the no-choice option does not have a substantive meaning but serves as an indicator of respondents' involvement with the task. In our applications we saw that the No-choice MNL model produced better results compared to the Nested MNL model. This may be an indication that in these applications the second explanation for choosing the No-choice may have been appropriate. In other words, when a conjoint choice experiment with a no-choice alternative is estimated with both the No-choice MNL model and the Nested MNL model, the (predictive) fit of the models may give an indication of the substantive reasons respondents have to choose the no-choice option. In particular, when the No-choice MNL model provides the best fit it may be inappropriate to interpret the estimates as reflecting the overall attractiveness of the category. It is interesting to note that in our second application the Nested MNL model did not provide any improvement in fit over the standard MNL model, while the No-choice MNL model did fit significantly better. We would like to note that in our applications the effect of the no-choice option itself was significant. If that is not the case, the model converges to the standard MNL solution and may neither fit nor predict better. However, the inclusion of the no-choice constant in the model allows for this test, and at the same time gives an indication of whether the standard MNL would be more appropriate.

References

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. In: B.N. Petrov & F. Csáki (eds) *2nd International Symposium on Information Theory*. Budapest: Akadémiai Kiadó, pp. 267–281.

Baron, J. & Ritov, I. (1994) Reference points and omission bias. *Organizational Behavior and Human Decision Processes*, 59, September, pp. 475–498.

Batsell, R.R. & Louviere, J.J. (1991) Experimental analysis of choice. *Marketing Letters*, 2, 3, pp. 199–214.

Ben-Akiva, M. & Lerman, S.R. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: The MIT Press.

Börsch-Supan, A. (1990) On the compatibility of Nested Logit models with utility maximization, *Journal of Econometrics*, 46, pp. 373–388.

Carroll, J.D. & Green, P.E. (1995) Psychometric methods in marketing research: Part 1, Conjoint analysis. *Journal of Marketing Research*, 32, 4, pp. 385–391.

Carson, R.T., Louviere, J.J., Anderson, D.A., Arabie, P., Bunch, D.S., Hensher, D.A., Johnson, R.M., Kuhfeld, W.F., Steinberg, D., Swait, J., Timmermans, H., & Wiley, J.B. (1994) Experimental analysis of choice. *Marketing Letters*, 5, 4, pp. 351–368.

Dhar, R. (1997) Consumer preference for a no-choice option. *Journal of Consumer Research*, 24, September, pp. 215–231.

Elrod, T., Louviere, J.J., & Davey, K.S. (1992) An empirical comparison of ratings-based and choice-based conjoint models. *Journal of Marketing Research*, 29, 3, pp. 368–377.

Huber, J. & Zwerina, K. (1996) The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, 33, 3, pp. 307–317.

Johnson, R.M. & Orme, B.K. (1996) *How Many Questions Should You Ask in Choice-Based Conjoint Studies*. Sequim, WA: Sawtooth Software.

Louviere, J.J. (1988) Conjoint analysis modelling of stated preferences. A review of theory, methods, recent developments and external validity. *Journal of Transport Economics and Policy*, January, pp. 93–119.

Louviere, J.J. & Woodworth, G. (1983) Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. *Journal of Marketing Research*, 20, 4, pp. 350–367.

McFadden, D. (1981) Econometric models of probabilistic choice. In C.F. Manski & D. McFadden (eds) *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: The MIT Press, pp. 198–272.

McFadden, D. (1976) Quantal choice analysis: a survey. *Annals of Economic and Social Measurement*, 5, 4, pp. 363–390.

Schwarz, G. (1978) Estimating the dimension of a model. *Annals of Statistics*, 6, pp. 461–464.