Apple Inc. v. Samsung Electronics Co. Ltd. et al

Doc. 991 Att. 19

Exhibit E

Dockets.Justia.com

11 Conjoint Choice Experiments: General Characteristics and Alternative Model Specifications

Rinus Haaijer and Michel Wedel

11.1 Introduction

Conjoint choice experimentation involves the design of product profiles on the basis of product attributes specified at certain levels, and requires respondents to repeatedly choose one alternative from different sets of profiles offered to them, instead of ranking or rating all profiles, as is usually done in various forms of classic metric conjoint studies. The Multinomial Logit (MNL) model has been the most frequently used model to analyze the 0/1 choice data arising from such conjoint choice experiments (e.g., Louviere and Woodworth 1983; Elrod, Louviere and Davey 1992). One of the first articles describing the potential advantages of a choice approach for conjoint analysis was by Madanski (1980). His conclusion was that conjoint analysts could adopt the random utility model approach to explain gross trends or predilections in decisions instead of each person's specific decision in each choice presented. The real breakthrough for conjoint choice came with the Louviere and Woodworth (1983) article in which they integrated the conjoint and discrete choice approaches.

The MNL model is the standard model for analyzing discrete choices, and can be derived from utility maximization (McFadden 1976). However, the MNL model does not accommodate heterogeneity of consumer choice behavior and potentially suffers from the Independence of Irrelevant Alternatives (IIA) property, which may be too restrictive in many practical situations. Latent class or mixture MNL models have been developed to accommodate heterogeneity (Kamakura, Wedel and Agrawal 1994). The Multinomial Probit (MNP) model does not suffer from IIA and deals with heterogeneity, but this model has some practical limitations related to identification, prediction and obtaining the choice probabilities. Haaijer et al. (1998) were the first to use a special specification of the MNP model for conjoint choice experiments.

In this chapter we review the alternative approaches to analyze conjoint choice experiments. But before doing that, we briefly describe in section 13.2 the general elements in conjoint analysis and the „classic" conjoint analysis approaches. Next, in section 13.3, the conjoint choice approach is discussed more extensively and an overview is given of recent conjoint choice applications in the marketing literature. Section 13.4 gives several approaches that can be used to estimate a conjoint choice experiment, including the MNL, the Latent Class MNL, and MNP models. These various models will be illustrated using an application to a conjoint choice experiment on coffee makers. Finally, section 13.5 compares the results of the various models and gives further discussion and conclusions.

11.2 General Concepts and Classic Conjoint Analysis

In marketing one wants to know which characteristics of products or services are important to consumers, for reasons of product optimization, new product design, price setting, market segmentation and competitive positioning amongst others. A technique, originally developed in the early 60's by Luce and Tukey (1964), that could eventually be applied to answer that question, is conjoint analysis. In conjoint analysis products or services are defined on a limited number of relevant attributes or characteristics each with a limited number of levels. These products, called profiles, have to be evaluated by respondents, who rank or rate them (as described in this section) or choose their most preferred ones from smaller choice sets (see section 13.3). As an introduction to conjoint choice experiments, in this section we describe briefly the general characteristics of conjoint analysis and the „classic” conjoint approaches, including ranking and rating conjoint. For a more extensive review see, e.g., Green and Srinivasan (1978, 1990), Louviere (1988) or Carroll and Green (1995).

The conjoint methodology is a decompositional approach to analyze consumer preferences. Product profiles are constructed from the product attributes, each defined at a certain number of levels, using factorial or fractional factorial designs (the latter to reduce the number of profiles and respondent burden in evaluating them). Respondents give an overall „score” to each product profile and the analyst has to find out what the preference contributions are for each separate attribute and level. Here it is commonly assumed that the overall utility of a profile is constructed by adding the preferences for the attribute-levels. This implies a compensatory preference model, in which a „low” score on a certain attribute can be compensated by a „high” score on another attribute. In conjoint experiments the contribution of an attribute (level) to the total utility is called a *„part-worth”*, and the total utility of a profile in a compensatory, additive preference model is equal to the sum of the part-worths: $U = \sum_s X_s \beta_s$, where U is the utility of the profile, X_s the value of attribute-level *s* and β_s is the weight parameter of attribute-level *s*. The part-worths can be computed from $X_s \beta_s$. More complex constructions are possible, such as a multiplicative model for the overall utility or interaction effects in the utility function.

Based on the analysis of the observed data several marketing questions can be answered (e.g., Vriens 1994) such as: 1) What is the (relative) importance of attributes and levels?, 2) What is the overall utility of specific profiles?, and 3) Are their individual differences?. Cattin and Wittink (1982) identified five different purposes for conjoint analysis in commercial applications: new product or concept identification, pricing, market segmentation, advertising and distribution. Later, competitive analysis and repositioning were added to this list (Wittink and Cattin 1989). Because conjoint analysis can be used for so many purposes, it has become a very popular marketing technique, with many applications in (commercial) marketing research (Cattin and Wittink 1982; Wittink and Cattin 1989; Wittink, Vriens and Burhenne 1994).

In a conjoint study several steps have to be taken. First of all, the attributes and the levels for each attribute have to be selected. Based on these attributes and levels the set of possible profiles can be constructed. However, it is easy to see that the total number of possible profiles can be very large even for a relative small number of attributes and levels. When there are for instance 3 attributes with 4 levels and 2 with 3 levels 576 ($4^3 \cdot 3^2$) different profiles can be constructed, which is clearly a too large number for respondents to rank or rate. Therefore, fractional factorial designs can be used to limit the total number of profiles in the analysis, while the main effects and first order interaction effects can still be estimated independently in many of these designs. The design one uses, and therefore the total number of profiles in the analysis, depends on how many interaction terms one wants to be able to estimate. In principle all kind of attributes, including price and brand, can be used in a conjoint study. However, the inclusion of brand as an attribute may lead to complications since it may represent implicit attributes such as quality (e.g., Oliphant et al. 1992; Struhl 1994). Having price as a separate attribute, orthogonal to the other attributes, may lead to unrealistic profiles, and care must be taken that no unrealistic price-brand, or price-attribute, combinations appear in the design. The selection of the number of levels of the attributes may also have some important implications. When all attributes have the same number of levels, the (absolute) values of the estimated part-worths give an indication of the (relative) importance of the attributes. However, it is not always possible to have the same number of levels for all attributes, since some attributes may be binary (e.g., a Yes/No or Present/Absent attribute) while others may have (many) more levels (e.g., „Brand”). Furthermore, Wittink et al. (1991) showed that when an attribute has more levels it becomes more important. They called this the „Number of Levels Effect”, an effect that has led to a substantial stream of research in its own.

Second, the evaluation task has to be selected. Above we mentioned ranking and rating tasks, but many more data collecting methods are available that all fall within the class of („classic”) conjoint analysis (see, e.g., Vriens (1995) for a detailed description of these methods), such as the full profile method (Green and Rao 1971), the tradeoff matrix method (Johnson 1974), the paired comparison method, Adaptive Conjoint Analysis (ACA) (Johnson 1985), or Hybrid Conjoint (Green, Goldberg and Montemayor 1981; Green 1984). All of these approaches can be used to obtain individual (segment or aggregate) level part-worths. Individual-level results are obtained using the observed „scores” of a respondent on the profiles and the characteristics of these profiles, and are often derived with regression-type procedures applied to each subject’s data. Subject characteristics or classification procedures may be used, however, for segmentation purposes, where respondents that perform similar on the conjoint task are put together in segments, which may be described using the subject characteristics.

Third, one has to choose the way the profiles are presented to the respondent and the way the data are collected (cf., e.g., Vriens 1995). The presentation of profiles can be done verbally, as a (printed) list of attributes and levels, with the use of pictorials, computer aided designs or actual products. Data collection can be done with a personal interview, a mailed questionnaire, over the telephone, or with a computer assisted procedure. Of course, some combinations of profile presentation and data collection are more suitable than others and some are not (always) possible. For instance, the construction of actual products is only possible for a very limited number of product categories because of the costs involved to actually produce all profiles in the experiment. See Vriens (1995) for a more extensive discussion on these issues.

Green and Srinivasan (1978) classified estimation methods for conjoint analysis in three categories. First, they described methods that assume that the dependent variable is, at most, ordinally scaled. In that case estimation methods like MONANOVA (Kruskal 1965), PREFMAP (Carroll 1972), or LINMAP (Srinivasan and Shocker 1973a/b; Pekelman and Sen 1974) can be used. Second, when it is assumed that the dependent variable is interval scaled, OLS regression techniques can be used. Third, for the paired comparison data in a choice context, the binary Logit or Probit model can be used. Tho se models arise as special cases of the models that we discuss more extensively later in this chapter.

In order to test the predictive ability of conjoint analysis, respondents most often have to evaluate a so-called holdout task after the main task. This task is usually similar to the main task, but the set of profiles differs. The responses on these holdout tasks are not used for estimation purposes but for prediction. The idea is of course that the estimated model should predict the holdout results as well as possible. Especially when no „real-life" data are available, the holdout task is simple way to test the predictive validity of a conjoint model. When no separate holdout task is obtained, predictive power can be tested by using the results of part of the respondents for estimation purposes to predict the results of the remaining respondents. However, this latter approach is only viable at the aggregate level.

The results of classic conjoint analyses are often used to predict choice or market share (Cattin and Wittink 1982). For instance, one may be interested to know what the predicted market shares of a specific product modification would be, or how the introduction of a new or modified product may affect the market shares of existing products in the market. To answer these kind of questions, market simulations have to be performed. In order to do this the individual level estimates have to be converted to choices to predict actual market behavior of the respondents. Many choice rules are possible, but one often-used method to achieve this employs the first-choice rule, where it is simply assumed that respondents choose the product with the highest utility. However, this approach may be inadequate because a deterministic rule is used to predict a probabilistic phenomenon (e.g., Louviere and Timmermans 1990). With the first-choice rule, the situation that an alternative has a probability of being selected over another alternative of 51% is treated the same as the situation that an alternative has a probability of 99% of being selected, which clearly present very different sets of preferences.

DeSarbo and Green (1984) listed five reasons why choice predictions constructed from the results of ranking or rating conjoint may not be accurate. They stated that (classic) conjoint studies are subject to incompleteness with respect to profiles, because the profile is never equal to the product, incompleteness with respect to model specification, because most often only main effects and some two-way interactions are estimated, and incompleteness with respect to situation, because conjoint assumes equal effects for marketing control variables across suppliers. Furthermore, they mentioned the artificiality of the conjoint analysis, caused by the fact that the amount of information in reality may be different from that in a conjoint experiment, and the instability of tastes and beliefs of consumers, because they may change over time. All of the above may be reasons that choice predictions are not accurate. However, DeSarbo and Green (1984) mention that aggregate market predictions from conjoint analysis can be quite good.

11.3 Conjoint Choice Experiments

11.3.1 Conceptual

Conjoint choice analysis has some advantages as compared to conventional conjoint analysis. There are no differences in response scales between individuals, choice tasks are more realistic than ranking or rating tasks, respondents can evaluate a larger number of profiles, choice probabilities can be directly estimated, and ad hoc and potentially incorrect assumptions to design choice simulators are avoided (Carroll and Green 1995). Several other authors point out similar (as well as some additional) advantages of the choice approach relative to the conventional approach (e.g., Louviere 1988; Elrod, Louviere and Davey 1992; Sawtooth Software Inc. 1995; DeSarbo, Ramaswamy and Cohen 1995; Cohen 1997; Vriens, Oppewal and Wedel 1998).

In the classic conjoint approaches described in the previous section, all profiles are presented to the respondent, while in the choice approach the total set of profiles is divided into several choice sets and respondents have to choose their most preferred alternative from each choice set. To set the scale of utilities between choice sets a base alternative often is added to each choice set. An advantage of the choice approach is that this base alternative not only can be one particular product profile, but it can also be a so-called „no-choice" option (see Haaijer, Kamakura and Wedel 2001 for a detailed discussion on the base alternative in conjoint choice experiments). In this case the choice probabilities can possibly be interpreted as market shares of the various profiles. The probability for the „no-choice" then might be interpreted as an indicator for the overall preference for the product category under research (e.g., Louviere and Woodworth 1983; Oppewal and Timmermans 1993). A disadvantage of including a no-choice alternative in the design is that respondents choosing that alternative provide no information on the alternatives and attributes and hence some information is lost (Elrod, Louviere and Davey 1992). Another potential problem with the no-choice option is the reason why respondents choose it. A reason could be that their preferred brand or price level is not in the choice set (or in general because of the presence -or absence- of a specific level of any attribute). Furthermore, a reason to choose the no-choice could be that respondents are not interested at all to do the task. Finally, they may find the choice too difficult and choose the no-choice if they decide not to spend more time on the choice task and avoid the difficult choice. In those cases one needs to be careful how to interpret the estimated no-choice probability. Johnson and Orme (1996) claim, after analyzing several conjoint choice experiments, that there is no evidence that the latter explanations may be true.

11.3.2 Design

The approach Louviere and Woodworth (1983) developed involved constructing conjoint choice experiments with the use of 2^J designs when there are *J* possible alternatives, obtained by generating all possible combinations of attribute levels. If there are, for instance, two attributes each with two levels, four alternatives can be constructed. The 2^J design used then contains all combinations of the four alternatives present or absent in the choice set. From the full 2^J design an orthogonal main effects experimental design is selected such that a relatively small number of choice sets remain for estimation purposes. A disadvantage of 2^J fractional factorial designs is that when there are many alternatives (*J*), this approach will result in large tasks for respondents where choice sets can contain (too) many alternatives. A more general version of the 2^J fractional factorial design can be used when each choice set contains a fixed number of alternatives (*M*) and each alternative has *S* attributes with each *L* levels. In that situation a $L^{M \cdot S}$ main effects, orthogonal, fractional factorial experimental design can be used to create joint combinations of attribute levels (e.g., Louviere and Woodworth 1983; Steenkamp 1985; Louviere and Timmermans 1990). In case the number of levels is not equal for all attributes a $L^{M \cdot S}$ design still can be used, where *L* now represents the maximum number of levels present in the study. Columns in the design representing attributes with fewer levels can be constructed by converting the columns with L levels to columns representing attributes with fewer levels.

The actual coding of levels in the choice designs can be done in several ways. For numerical attributes (e.g., price) actual values can be used in the design, which leads to so-called linear attributes. However, most of the time some dummy specification is used. This specification can involve „regular" dummy coding (e.g., „1" if a level is present and „0" if it is not present) or so-called effects-type coding. In the situation of 3 levels of an alternative, with effects-type coding, the first level is coded, e.g., as $[1\,0]$, the second as $[0\,1]$ and the third as $[-1\,-1]$. For attributes with 2 levels the codes are +1 and -1 respectively. This way of coding has as advantage, when all attributes are coded this way and each level appears with equal frequency in the design, that the sum of the part-worths for each level is equal to zero, so that the total model is centered around zero. Combinations of different ways of coding are possible.

A specific characteristic of conjoint choice experiments is that one needs two designs, instead of one design in the classic conjoint approach, to set up the experiment. One design is needed to construct the profiles, like in the classic conjoint approach, but an additional design is needed to put these profiles in various choice sets. It is beyond the scope of this chapter to discuss extensively how efficient designs for conjoint (choice) experiments should be constructed, but the key elements are described briefly. For much more detail the interested reader is referred to, e.g., Addelman (1962), Louviere and Woodworth (1983), Steenkamp (1985, in Dutch), Kuhfeld, Tobias and Garratt (1994) or Huber and Zwerina (1996). In principle one wants the main effects and interaction effects to be orthogonal in the design, however, Kuhfeld, Tobias and Garratt (1994) show that orthogonal designs are not always more efficient than non-orthogonal designs, hence a trade off has to be made between these two concepts. Furthermore, they show that the efficiency of a given design is affected by the coding of quantitative factors, even though the relative efficiency of competing designs is unaffected by coding (Kuhfeld, Tobias and Garratt 1994, p. 549).

The range of levels for quantitative factors should be as large as possible to maximize efficiency. However, the levels should of course not be implausible. Huber and Zwerina (1996) describe four properties that characterize efficient choice designs. They mention level balance, orthogonality, minimal overlap and utility balance. Level balance means that each level of an attribute appears with equal frequency. However, level balance and orthogonality are often conflicting. Choice sets should have minimal overlap since alternatives that have the same level of an alternative provide no information on the preference for that attribute. Hence, the probability that an attribute level repeats itself in each choice set should be as low as possible. Level balance, orthogonality and minimal overlap are used to construct optimal utility-neutral designs. The efficiency of such design can be improved by balancing the utilities of the alternatives in each choice set. This is important since choice sets that generate extreme probabilities are less effective at constraining the parameters of the choice model than are moderate ones (Huber and Zwerina 1996, p. 308), although they do have a big positive impact on the log-likelihood of a choice model. So, a high likelihood may go together with imprecise parameter estimates for choice sets with more extreme probabilities. One possible way to achieve more utility balanced designs is simply by re-labeling the levels of the attributes, which has as advantage that it does not affect orthogonality, in contrast to swapping techniques. One problem not solved yet is how efficient designs can be obtained when a base alternative (such as a no-choice) is present in the experiment (Huber and Zwerina 1996), another is that efficient designs for the MNP model have not been developed yet, although recently design procedures for the related mixed logit model have been proposed (Sándor and Wedel 1999).

Another issue that plays a role is the type of design to use in the analysis: a design with fixed, randomized or individualized choice sets. With a fixed choice set approach each respondent (or each group of respondents, in a slightly more general fixed approach) receives exactly the same choice sets at exactly the same stage of the choice task. In a randomized experiment each respondent (or group of respondents) also receives the same choice sets but in a different order to compensate for learning and fatigue effects that are expected to average out in this way. In an individualized experiment each respondent receives his own choice sets. An advantage of individualized choice sets is that it can be tested whether preferences, or attribute importance, change in later stages of the choice experiment (Johnson and Orme 1996), since for each choice set (i.e., the 1st, 2nd, ..., last for each respondent) estimates can be obtained in this situation. A study by Johnson and Orme (1996) showed, when comparing several conjoint choice experiments, that the importance of brand decreases throughout a conjoint choice experiment, while that of price increases. A disadvantage of using individualized choice sets is that no choice frequencies can be computed for alternatives in each choice set, since each set is only evaluated by one respondent. Another disadvantage of individualized choice sets is that comparison and clustering becomes more difficult (Oliphant et al. 1992), which would be possible when all (groups of) respondents receive the same choice sets. In that latter case, respondents that show similar choice patterns can be grouped together in segments. This problem is however alleviated by mixture model approaches to conjoint choice experiments, as detailed below. Depending on what type of analysis, the fixed, randomized or individualized approach can be the preferred choice.

11.3.3 Applications

In this section an overview is given of recent applications of conjoint choice experiments in the marketing literature. Some of the studies listed here have been already discussed briefly in the previous section. This overview is not intended to be complete, but the aim is to give an impression of possible applications of conjoint choice. In particular, we will show for each study several characteristics of the conjoint experiment. Table 1 lists the studies we present in this overview (it was not possible to retrieve all information for all studies) .

Table 1 shows that the range of products investigated in conjoint choice experiments is rather wide. The products range from fast-moving consumer goods, like toothpaste, to durable products, like houses and cars. The same holds for the number of choice sets presented to the respondents and the number of alternatives in the choice sets. In the various applications, respondents had to choose from 3 to 32 choice sets containing 2 to 8 alternatives. The profiles in these choice sets were defined on 2-12 attributes. In the Oppewal, Louviere and Timmermans (1994) study 33 attributes where used, but their aim was to reduce this number using Hierarchical Information Integration.

The number of respondents used in the various studies also shows a wide range from 64 up to almost 1000 respondents. There seems to be more agreement about the type of base alternatives to use in a conjoint choice study. Most of the studies listed in Table 1 used „none", „own" or „other" base alternatives and only a few used a fixed profile as base. Most of the studies that did use a fixed base alternative assume a specific situation (for instance like „given that you are going on a holiday, what would be your most preferred trip") and are less interested in obtaining market shares, which is the major advantage of including some sort of „no-choice" base alternative. This may be the main reason to include such a base alternative in the other applications. Note that for the studies with only two alternatives no base alternative was used. The number of levels of the attributes used in the studies also shows a rather consistent pattern. Most studies use attributes with 2-4 levels. In situations that more levels are used for an attribute, this most often is a brand-attribute.

Table 1: Conjoint Choice Applications

Authors	Product/ Product category	Attri-butes	Choice sets	Alter-natives*	Respon-dents	Base	Levels / Design
Elrod, Louviere and Davey (1992)	Rental apartments	4	27	3	115	Own	$2^4 \cdot 3^4$
Oliphant et al. (1992)	Insurance	9	20	5	149	None	$4^2 \cdot 2^8$
Oppewal and Timmermans (1992)	Shopping centers	4	8/16	3 /4	?	Other	4^4
Chrzan (1994)	Mail orders	5	8	3	605	None	2^5
	Fashion access.	?	16	6	300	Other	$8 \cdot 4^2 \cdot 3^4$
	Consumer fashion	10	16	3	876	Other	$4 \cdot 3 \cdot 2^8$
Oppewal, Louviere and Timmermans	Shopping centers	33	3	3 /4	396	None	$4^5 \cdot 2^3 + 4^8 \cdot 2^5$
(1994)	Batteries	3	12/24	2	65	-	3^3
Allenby, Arora and Ginter (1995)	Credit cards	7	13-17	2	946	-	$4^2 \cdot 3^3 \cdot 2^2$
Allenby and Ginter (1995)							
	Activity packages	4	5/6	3	221	Fixed	3^4
Dellaert, Borgers and Timmermans (1995)	Food	2	16	8	600	Fixed	?
DeSarbo, Ramas-wamy and Cohen (1995)	Houses	4	16	3	278	None	4^4 (4^8)
Timmermans and Van Noortwijk	Flower exhibits	3	?	3	64	Fixed	$3 \cdot 2^2$
(1995)	Tourist Portfolio	12	12	3	±660	Fixed	3^{12}
Dellaert, Borgers and Timmermans (1996)	Toothpaste	5	32	5	184	None	?
Dellaert, Borgers and Timmermans (1997)	Coffee makers	5	4/8	5/3	185	Fixed	$3^3 \cdot 2^2$
Moore, Gray-Lee and Louviere (1998)	Cars	6	9	4	200	None	$8 \cdot 3^3 \cdot 2^2$
Vriens, Oppewal and Wedel (1998)							
Wedel et al. (1998)							

*: Base included

The information in Table 1 shows that conjoint choice experiments can be and have been used for a wide range of possible applications. In almost any situation in which consumers have to choose between several options the conjoint approach can be used to determine which attributes of the product are important for the respondent. In this case „product” can be some fast moving product like toothpaste, a durable like a car or a house, a service such as tourist attractions, and even products that are not actually bought by respondents, like „shopping centers”. In most of the recent applications the number of attributes, levels and choice alternatives used in the design is rather low (attributes and alternatives around 4, levels around 3), although there are some exceptions. The number of choice sets that can be presented to respondents showed a wide range. A recent study by Sawtooth Software (Johnson and Orme 1996) showed that given the rather short response times in conjoint choice experiments, many choice sets can be offered to respondents even without reducing the quality of the choices. With modern computer assisted data collecting methods for conjoint choice the response times can be obtained very easily, and can actually be used to improve estimation of partworths, see Haaijer, Kamakura and Wedel (2000).

11.4 Conjoint Choice Models

11.4.1 Introduction

In this section we discuss several approaches to analyze conjoint choice experiments. First of all, the standard MNL approach is discussed in section 13.4.3. Second, section 13.4.4 describes the Latent Class MNL model. Section 13.4.5 provides two MNP models, one in which choice sets are assumed independent, and one where the choices from one individual are treated as correlated. But first we specify the general structure of conjoint choice models in this section and describe the data we will use as application in section 13.4.2.

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 set of profiles over K choice sets. In this chapter we assume that each choice set contains the same number of alternatives, without losing generality. In order to formulate models for conjoint choice experiments, we start from random utility maximization (McFadden 1976). The utility of alternative m in choice set k for individual j is defined as:

(1) $$U_{jkm} = X_{km}\beta + e_{jkm},$$

where X_{km} is a (1xS) vector of variables representing characteristics of the mth choice alternative in choice set k, β is a (Sx1) vector of unknown parameters, and e_{jkm} is the error term. Note that we assume that the X-matrix in (1) does not depend on j, because in conjoint choice experiments no individual characteristics appear in the analysis in general. Note, however, that when an individualized design is used X does depend on j, but we omit this index here for convenience.

For each individual *j*, it is assumed that the alternative with the highest utility is chosen and a variable y_{jkm} is observed which is for each choice set *k* defined as:

$$(2) \qquad y_{jkm} = \begin{cases} 1 \text{ when } U_{jkm} > U_{jkn} \ \forall \ n \neq m \\ 0 \text{ when } \exists \ n \neq m : U_{jkn} > U_{jkm} \end{cases}, \ n = 1,..,M \ .$$

As mentioned in section 13.3.1, in conjoint choice experiments a base alternative is often used in each choice set *k* to scale the utility over choice sets. This base alternative cannot only be a regular profile, it also can be specified as a no-choice alternative („None of the above”) or an „own-choice” alternative („I keep my own product”). This kind of base alternative, however, presents the problems of how to include it in the design of the choice experiment, and in what way to accommodate it in the choice model. Regular choice alternatives are most often coded in the design matrix with effect-type or dummy coding. Since the no-choice alternative does not possess any of the attributes in the design, it is often coded simply as a series of zero’s, which makes the fixed part of its utility zero in each choice set. However, the utility level of the no-choice alternative still has to be specified when effect-type coding is used, since the zeros of the no-choice act as real levels in that case and this potentially leads to biased estimates. The no-choice alternative can be specified in two ways. The first specification is to include a no-choice constant in the design matrix *X* in (1). This introduces one additional parameter in the model to estimate. Note that when brand dummies (or other attribute specific dummies) are used for each level of the attribute, no additional parameter is needed since in that case the utility level of the no-choice is already set by those dummies. However, the total number of parameters to estimate is equal in both cases. The second specification is to formulate a nested MNL model, in which it is assumed that subjects first choose between the no-choice and the other choice alternatives in the choice set, and in a second stage make their choice among the alternatives when they decide to make a „real” choice. This also introduces one additional parameter in the model, the *dissimilarity coefficient* of the Nested MNL model. Which of these representations for the no-choice option is preferable is discussed in Haaijer, Kamakura and Wedel (2001).

11.4.2 The Data

The various models that will be introduced in the next sections to analyze the above conjoint choice structure will be illustrated with an application, which is a replication of part of an application reported by Haaijer et al. (1998), with coffee-makers as the product category. The five attributes, and their levels, for the coffee-makers are listed in Table 2. Using a factorial design, sixteen profiles were constructed. Data were collected from 185 respondents, divided into two groups that received different choice sets based on the same sixteen profiles. Respondents had to choose from eight sets of three alternatives. Each choice set included the same base alternative, which is a fixed regular alternative in this experiment. Furthermore, eight holdout profiles were constructed, which were divided into four holdout sets with three alternatives, where the same base alternative was used as in the estimation data. These holdout sets were offered to all respondents. The estimation and holdout designs were coded using effects-type coding.

For all models in subsequent sections we will obtain parameter estimates. Furthermore, to compare model performance, we report the log-likelihood value, AIC and BIC statistics, and Pseudo R^2 value (e.g., McFadden 1976) relative to a null-model in which the probabilities in a choice set are equal for all alternatives. The AIC criterium (Akaike 1973) is defined as: $AIC = -2 \ln L + 2n$ and the BIC criterium (Schwarz 1978) is defined as: $BIC = -2 \ln L + n \ln(O)$, where $\ln L$ is the log-likelihood in the optimum, n the total number of estimated parameters in the model, and O the number of independent observations in the experiment.

Table 2: Attributes and Levels of Coffee-Makers

Attribute Level	Brand	Capacity	Price (Dfl)	Special Filter	Thermos-flask
1	Philips	6 cups	39,-	Yes	Yes
2	Braun	10 cups	69,-	No	No
3	Moulinex	15 cups	99,-		

11.4.3 Multinomial Logit

The most popular discrete choice model is the Multinomial Logit (MNL) model. It follows when the assumption is made that the error term in (1), e_{jkm}, is independently and identically distributed with a Weibull density function. A Weibull density function for a random variable Y is defined as (see, e.g., McFadden 1976):

$$(3) \qquad P(Y \leq y) = \exp^{-\exp^{-y}}.$$

This distribution belongs to the class of double negative exponential distributions as are, e.g., the Type I extreme value distribution and the Gumbell distribution, which are sometimes also used to specify the MNL model. The MNL model treats observations coming from different choice sets for the same respondent as independent observations. Therefore, in estimating the MNL model, 100 respondents choosing from 10 choice sets yields the same computational burden as 1000 respondents choosing from 1 choice set. In the standard MNL model, with one choice observation for each individual, the choice probabilities have a simple closed form. The choice probabilities in the conjoint MNL approach can be obtained through a straightforward generalization of this standard model.

The probability p_{km} that alternative *m* is chosen from set *k* is in this case simply equal to (cf., e.g., Maddala 1983, p. 60-61; Ben-Akiva and Lerman 1985; Swait and Louviere 1993):

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

The standard log-Likelihood for discrete choice models is in the conjoint context for the MNL model extended by adding a summation over choice sets:

(5) $$L_{MNL} = \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{m=1}^{M} y_{jkm}\ \ln\ (p_{km}).$$

The simple and easy to calculate form of the choice probabilities (4) in the MNL model has much contributed to its popularity in conjoint choice experiments. However, there is a serious limitation to the use of this model that is related to the *Independence of Irrelevant Alternatives* (IIA) property. This property arises from the assumption of independent random errors and equal variances for the choice alternatives, and implies that the odds of choosing one alternative over another alternative must be constant regardless of whatever other alternatives are present (e.g., Louviere and Woodworth 1983; Ben-Akiva and Lerman 1985), which may often be too restrictive in practical situations. If it is assumed that the IIA property holds and the MNL model is used, predicting the choice probabilities of new alternatives can simply be done by inserting the attribute values of these new alternatives in the closed form expressions for the choice probabilities (4). Green and Srinivasan (1978) stated that in consumer behavior contexts the IIA property might not be a realistic assumption, especially when some of the alternatives are close substitutes (cf. McFadden 1976). When covariances across alternatives are incorrectly assumed to be zero, the estimates for the effects of explanatory variables are inconsistent (Hausman and Wise 1978; Chintagunta 1992). When the IIA property does not hold, other models that avoid IIA, should be used instead of the standard MNL model, however, at the cost of computational complexity. One of the most general of these models is the Multinomial Probit (MNP) model, which is discussed in section 13.4.4.

When the IIA assumption is true, the parameters of the Logit model can be estimated when the sufficient condition is satisfied that the alternatives are independent across choice sets (Louviere and Woodworth 1983). So, choices between alternatives must be pairwise independent across choice sets. The alternatives in a conjoint choice experiment are obtained by using an orthogonal, fractional factorial main effects design (Louviere and Woodworth 1983; Louviere and Timmermans 1990). A constant base alternative is useful, because it preserves the design orthogonality of the attribute vectors of conjoint alternatives (Louviere 1988;

Elrod, Louviere and Davey 1992). However, in the case of the Logit model, design orthogonality does not imply information orthogonality, for which the parameters would be uncorrelated. When similarities across alternatives are incorrectly assumed to be zero, the estimates for the effects of marketing variables are incorrect (e.g., Chintagunta 1992).

The expression for the choice probabilities (4) may be expanded to accommodate ranking data, which is particularly useful in conjoint analysis (McFadden 1986; Kamakura, Wedel and Agrawal 1994). However, the assumptions needed to translate rankings into choices may not hold, especially when individuals use elimination and nesting strategies the IIA property does not hold (Louviere 1988). Also, the use of brand names in the conjoint design may result in correlations between the utilities of the alternatives, violating the IIA property. In order to be able to test for IIA, design plans that allow as many relevant two-way interactions as possible to be tested can be used (Louviere and Woodworth 1983).

Table 3: *MNL Estimation Results*

Attribute (level)	*Est.*	*S.e.*
β_1 Brand (1)	0.040*	.013
β_2 Brand (2)	-0.329*	.013
β_3 Capacity (1)	-1.015*	.014
β_4 Capacity (2)	0.494*	.010
β_5 Price (1)	0.313*	.018
β_6 Price (2)	0.372*	.013
β_7 Filter (1)	0.340*	.070
β_8 Thermos (1)	0.312*	.010
Statistics		
Ln-likelihood	-1298.706	
AIC	2613.412	
BIC	2655.810	
Pseudo R^2	0.201	

*: $p < 0.05$.

The MNL specification was used to analyze the data set described in section 13.4.2. In Table 3 the parameter estimates and fit statistics are listed. With the effects-type coding used, the part-worth for the last level of each attribute can be constructed by taking the sum of the estimates of the other levels of that attribute and change the sign. The results show that respondents prefer a high capacity to a low capacity, a low price level over a high price level, and that they prefer the presence of a special filter and thermos flask to the absence of those attributes. Finally, the third brand is the most attractive brand and the second the least. The pseudo R^2 has a value of 0.201, which for this kind of choice data is a reasonable value. The estimates in Table 3 were used to predict the holdout sets. This resulted in a predicted log-likelihood of -754.860 (Pseudo R^2 =0.072). This shows that the MNL model does not a very good job in predicting the holdout sets in this application, since the obtained Pseudo R^2 has a value much lower as those resulting from the estimation sample.

The standard MNL model described in this section assumes that all respondents act similar in their choice behavior. However, several groups of respondents may exist that show different choice behavior. The next section describes the Latent Class MNL model that can be used to obtain segments of respondents.

11.4.4 Latent Class MNL

Next to its disadvantages related to the IIA assumption, the MNL model also suffers from the problem that it treats all subjects in the sample as homogeneous, and does not deal with heterogeneity. The MNL model cannot be estimated at the individual level, and thus subject-specific part-worths cannot be obtained (e.g., Elrod, Louviere and Davey 1992). The issue of subject heterogeneity has received a lot of attention in the marketing literature and has become a topic of much research (cf. Wedel et al. 1999). Basically, there are two ways to accommodate heterogeneity. In this section we deal with one, where one specifies a discrete distribution of the response coefficients β_j across the population, that is, one postulates that groups of respondents exist with different part-worths. This leads to latent class or finite mixture discrete choice models, which have been applied to conjoint choice experiments by Kamakura, Wedel and Agrawal (1994), and DeSarbo, Ramaswamy and Cohen (1995). In the subsequent sections we deal with Multinomial Probit models that specify a continuous distribution of heterogeneity. Finite mixture models connect very well to marketing theories of market segmentation (Wedel and Kamakura 1997) and have enjoyed considerable success. Managers seem comfortable with the idea of market segments, and the models appear to do a good job of identifying segments from conjoint choice data.

Kamakura, Wedel and Agrawal (1994) developed a unifying mixture regression model for segmentation of choice data. Their approach assumes that choices are based on random utility maximization. The observed choice variables y_{jkm}, are assumed to be independent multinomial, and to arise from a population that is a mixture of Q unobserved segments, in proportions $\pi_1, \ldots, \pi_Q$. We do not know in advance from which segment a particular subject arises. The probabilities π_q are subject to the following constraints.

(6) $$\sum_{q=1}^{Q} \pi_q = 1 \quad \pi_q \geq 0 \quad q = 1,\ldots,Q$$

Given segment q, the choice probability for profile m for choice set k is:

(7) $$P_{km|q} = Prob[U_{km|q} \geq U_{kn|q} \;\; n = 1,\ldots,M, n \neq m],$$

where $U_{km|q}$ is the random utility derived from alternative *m* at *k* in segment *s*. Consumers are assumed to maximize their utility over the entire choice set, $U_{km|q} = \max\{ U_{kn|q} \; n = 1, \ldots M \}$. As before, the random utility for segment *q* is assumed to be a function of the attributes:

(8) $$U_{km|q} = X_{km} \; \beta_q + \varepsilon_{kmq} \, .$$

If the random components, ε_{kmq}, are assumed to be independent and identically Weibull distributed, the choice probabilities for segment *q* are:

(9) $$P_{km|q} = \frac{\exp[U_{km|q}]}{\sum_{n=1}^{M} \exp[U_{kn|q}]} \, .$$

The idea behind the mixture model is that if the probability conditional upon knowing the segments have been formulated, the unconditional probability of observing the *K* choices is obtained as:

(10) $$P_j = \sum_{q=1}^{Q} \pi_q \prod_{k=1}^{K} \prod_{m=1}^{M} P_{km|q}^{Y_{jkm}} \, .$$

As additional features of the model, the prior probabilities of segment membership can be reparameterized according to a concomitant variables model and the model can deal with rank-ordered data (Kamakura, Wedel and Agrawal 1994). Wedel et al. (1998) extended this mixture model for conjoint choice experiments by assuming that the brand can be decomposed into latent dimensions and segment-specific ideal points along those dimensions. The utility function further contains a linear combination of the attribute-level dummies. This model integrates conjoint analysis and multidimensional scaling, which makes it especially suited for product positioning.

The mixture regression model for conjoint choice experiments was applied to the coffee-maker data for $Q = 1$ up to 6 segments; the BIC statistic indicated $Q = 4$ segments as optimal. The four- segment solution was run ten times from different starts to overcome problems of local optima. Table 4 gives the estimated parameters for all 4 segments.

Table 4 shows that the four segments differ in their preferences for the attributes. The first segment (49,7% of the sample) wants a coffee-machine that contains as many features as possible, for a as low price as possible, but it does not matter what brand it is, since the brand parameters are not significant different from zero for this segment. So, this seems to be a price-sensitive segment. The second segment (17,4%) does not want a low capacity machine and prefers one with a thermos-flask, but this segment seems more quality seeking, since its partial utility for the lowest price level is negative, while that for the highest price level has the highest partial utility. This segment also has no preference for a particular brand. The third segment (13,5%) also likes only one feature (a special filter in this case), but does not want to pay a high price for it. This segment in addition has a preference for one of the brands (brand 2). Finally, the fourth segment (19,4%) finds the brand of the coffee-machine most important. It has a high preference for the first and third brand, and in addition for the presence of a special filter. So, this segment seems to be a real brand-seeking segment.

Table 4: LCMNL Estimation Results

Attribute (level)	*Segm. 1*		*Segm. 2*		*Segm. 3*		*Segm. 4*	
	Est.	*S.e.*	*Est.*	*S.e.*	*Est.*	*S.e.*	*Est.*	*S.e.*
β_1 Brand (1)	0.005	.179	0.289	.188	-0.051	.316	0.920*	.202
β_2 Brand (2)	-0.224	.163	-0.321	.179	0.683*	.295	-1.673*	.346
β_3 Capacity (1)	-2.681*	.289	-0.956*	.247	-1.191*	.481	-0.001	.165
β_4 Capacity (2)	1.302*	.152	0.627*	.175	0.578*	.266	0.355	.203
β_5 Price (1)	1.263*	.284	-0.934*	.346	1.431*	.359	-0.322	.271
β_6 Price (2)	0.698*	.213	0.150	.193	0.605	.312	-0.064	.218
β_7 Filter (1)	0.630*	.132	0.134	.133	1.753*	.340	0.390*	.133
β_8 Thermos (1)	0.461*	.139	0.876*	.167	0.289	.159	-0.053	.167
Segment Size	0.497		0.174		0.135		0.194	
Statistics								
Ln-likelihood	-1040.271							
AIC	2115.542							
BIC	2336.058							
Pseudo R^2	0.360							

*: $p < 0.05$.

From the results of an Latent Class analysis different marketing strategies can be developed for the various segments, provided that these are big enough to make it profitable to develop a specific strategy. Table 4 shows that the smallest segment (segment 3) still contains 13.5% of the respondents, which may be big enough to target. The estimates in Table 4 were used to predict the likelihood of the holdout choice sets. This gives a predicted log-likelihood for the LCMNL model of -708.832 (Pseudo R^2 =0.128). Comparing this to the predictive fit of the MNL model (Pseudo R^2 =0.072) we see a substantial improvement. Thus the LCMNL model improves upon the MNL model by accommodating heterogeneity, by providing actionable information on market segments, and by providing better holdout predictive performance.

Although the LCMNL model accounts for consumer heterogeneity, it still treats choices made by the same respondent as independent. In the next section two versions of the MNP model will be developed. One in which it is still assumed that choice sets are independent, but choice alternatives within a choice set may be correlated, and one that in addition relaxes the independence of choice sets.

11.4.5 Multinomial Probit

Introduction

In the previous section we saw how latent class MNL model can be used to account for heterogeneity. The MNP model also presents a way to deal with continuous heterogeneity distributions of the part-worths across consumers. Specifically, the parameters are specified to follow a normal distribution, which by some has been argued to be a better representation of heterogeneity than a discrete mixing distribution (Allenby and Rossi 1999). The continuous heterogeneity distribution has several advantages: It has been argued to characterize the tails of heterogeneity distributions better and may predict individual choice behavior more accurately than finite mixture models, since the tails may have a substantial impact on the predictive performance of the models. It provides a parsimonious representation of heterogeneity and flexibility with regard to the appropriate choice of the distribution of heterogeneity (see, e.g., Arora, Allenby and Ginter 1998).

In the conjoint choice context, the Multinomial Probit model (MNP) offers the major advantage of allowing correlations among the repeated choices that consumers make from the multiple choice sets next to allowing correlation of random utilities of alternatives within choice sets. This follows since the MNP model relaxes the assumption of independence of the error terms in random utility models (e.g., Daganzo 1979; Kamakura 1989), and thereby alleviates IIA. Factors such as learning, boredom, or anchoring to earlier choice tasks may distort the measurement of preferences when these are assumed independent, like in the MNL model, and these effects should be tested and/or accounted for (McFadden 1986). Several studies showed that subjects' utilities for alternatives may indeed depend on the choice context (e.g., Huber, Payne and Puto 1992, Simonson and Tversky 1992; Nowlis and Simonson 1997), where „context" is defined as the particular set of alternatives evaluated. Since the design of conjoint choice analysis involves only a subset of all possible profiles (constructed by fractional factorial designs) and choice sets that vary in composition (constructed by blocking designs), context effects are likely to occur in those experiments. Simonson and Tversky (1992) distinguished *local contrast effects* and *background contrast effects*. Local contrast effects are caused by the alternatives in the offered set only, while background contrast effects are due to the influence of alternatives previously considered. In a conjoint choice experiment, local contrast effects may occur due to the composition of a particular choice set in terms of the attribute levels of the profiles, affecting attribute importance, inducing correlations among the utilities of profiles in the choice set and leading to a violation of IIA (Simonson and Tversky 1992). On the other hand, background contrast effects may occur in conjoint choice experiments if the attribute importance of profiles in a particular choice set are influenced by tradeoffs among profiles in previous choice sets. In this case covariance among the random utilities of alternatives in different choice sets may occur. This violates the assumption of independence of choices among alternatives in different sets, as assumed in the MNL model. For a more extensive discussion on context effects see, e.g., Tversky (1972), Huber, Payne and Puto (1982), Huber and Puto (1983), Simonson (1989), or Simonson and Tversky (1992).

Timmermans and Van Noortwijk (1995) explicitly modeled context effects in conjoint choice experiments by including cross effects in the design matrix such that the utility of an alternative depends on its own and other alternatives' attributes. However, they only model context effect within choice sets (i.e., the local context effects) and not between choice sets (i.e., the background context effects). Haaijer et al. (1998) applied the MNP model to conjoint choice experiments, accounting for both the local and the background context effects described above through a specific covariance structure. They showed that it is important to account for both types of context effects. Two sections below give a MNP model that deals with heterogeneity, IIA and local context effects, and a MNP model that in addition accounts for background context effects, respectively.

First we specify the MNP model in general, starting again from the utility function. Assume again that there are J respondents, each receiving the same H profiles which are divided into K smaller sets with M alternatives each. A base alternative that is common to all sets is added to the profiles and scales the utility levels between choice sets. This base alternative can be a no-choice alternative or a regular profile. The other profiles are unique to their particular choice set, so that $H = K(M1)+1$. The utilities of the alternatives for individual j are contained in the latent unobservable vector u_j, which satisfies:

(11) $$u_j = X\beta_j + e_j,$$

where X is a (HxS)-matrix containing the attributes of the alternatives, β_j is a (Sx1) vector of random weights, and e_j is the vector containing the random component of the utilities.
In the MNP model is it assumed that e_j is distributed as:

(12) $$e_j \sim N_H\ (0, \Sigma_e),$$

independent between individuals; Σ_e is a (HxH) positive definite covariance matrix. In the MNP model not only the -parameters in (11) have to be estimated but also the parameters in the covariance matrix Σ_e.

A potential problem of the MNP model is that of identification. A model is identified when there is only one set of estimates that maximizes the likelihood. When different parameter estimates give the same results, interpretation of the estimates becomes difficult. Bunch and Kitamura (1989) demonstrated that nearly half of the published applications of MNP are based on non-identified models. It is easy to see that when the covariance matrix in (12) of the MNP model is multiplied with a factor β and all -estimates in (11) with a factor $\sqrt{\alpha}$, that this leads to the same results. So, at least one parameter in the MNP model must be fixed to scale the model and to identify the other parameters. Often, one of the variance parameters is used for this purpose, but this is not sufficient, however. In the standard MNP model with one choice set (*K*=1), only $M(M-1)/2-1$ of the $M(M+1)/2$ covariance parameters in Ω are identified (Dansie 1985; Bunch 1991; Keane 1992). So, $M+1$ restrictions must be imposed on the Ω-matrix in this situation.

Furthermore, in conjoint choice experiments, (holdout) predictions are often required. Another problem, besides identification, of the general MNP formulation is that these predictions for new profiles, not included in the conjoint design, cannot be made with the covariance matrix in (12) because in predicting choice probabilities for alternatives not included in the design of the experiment, estimates of the covariances of these new profiles are required and those are not available (cf., e.g., Pudney 1989 p.115; Elrod and Keane 1995; Haaijer et al. 1998).

In order to arrive at an MNP model for conjoint choice experiments that is both identified and that allows for predictions of new profiles, restrictions have to be imposed on the covariance matrix. We allow for heterogeneity in the attribute level coefficients by specifying β_j in equation (11) as (cf., e.g., Hausman and Wise 1978; Daganzo 1979; Ben-Akiva and Lerman 1985):

(13) $$\beta_j = \beta + \psi_j ,$$

with $\psi_j \sim N_S(0, \Sigma_\psi)$, independent of e_j. Then

(14) $$u_j \sim N_H(X\beta, \Omega),$$

with:

(15) $$\Omega = \Sigma_e + X \Sigma_\psi X' .$$

The specification that enables the prediction of new alternatives that we use assumes $\Sigma_e = I_H$, and for reasons of parsimony and identification we parameterize Σ_ψ as a matrix of rank one: $\Sigma_\psi = \sigma \sigma'$, with an *S*-vector of parameters, where *S* is the number of columns in the *X*-matrix. The number of parameters in Ω now is equal to the number of β-parameters. Especially when the number of columns (*S*) in *X* or the number of profiles (*H*) is large, this specification for Ω is very parsimonious compared to a full random coefficients model or general Probit model. A more general specification for Ω results in an increase in the number of covariance parameters so that identification often becomes a problem. So, we now have:

(16) $$\Omega = I_H + X \sigma \sigma' X' .$$

This random coefficients model may account for heterogeneity, violations of IIA, and local and background context effects potentially caused by all attributes in the conjoint design. See Haaijer et al. (1998) and Haaijer (1999) for a more extensive discussion of this specification and its characteristics. Rossi, McCulloch and Allenby (1996) developed a related random coefficients Bayesian MNL model.

Like for the MNL and LCMNL model, estimates for the parameters are obtained for the MNP model by maximization of the likelihood (see below) over β and the parameters in the covariance matrix. However, when there are more than three alternatives in a choice set the choice probabilities cannot be evaluated numerically in the MNP model (cf., e.g., McFadden 1976; Maddala 1983; Kamakura 1989; Keane 1992). Simulation techniques have been developed that solve this problem of the MNP model. To obtain the estimates in the MNP models in the next two subsections, the Simulated Maximum Likelihood (SML) method is applied using the SRC simulator. A discussion of simulation techniques is beyond the scope of this chapter, for an extensive discussion see, e.g., Hajivassiliou (1993).

Multinomial Probit with independent choice sets

A straightforward way to apply the MNP model is to use it in the similar way as the MNL model. In this case we take an individual's utilities to be independent between the choice sets, and thus account for local, but not for background context effects. We then have *JK* independent observations, and the log-likelihood is again a straightforward generalization of the standard likelihood of choice models, where a summation over choice sets is introduced, similar as in the MNL model of section 13.4.3 Letting p_{km} denote the fraction of individuals choosing alternative *m* in set *k*, the log-likelihood is (in a slightly different notation as the log-likelihood (5) of the MNL model) equal to:

$$(17) \qquad L_{MNP_{csi}} = J \sum_{k=1}^{K} \sum_{m=1}^{M} p_{km} \ln (\pi_{km}),$$

where π_{km} is the probability that alternative *m* is chosen in set *k*. Note again that in conjoint choice models consumer characteristics or other individual specific variables are usually not included, hence π_{km} does not depend on *j* and each individual has the same probability of choosing any specific alternative, since we assumed that they all receive the same choice sets. This model is called the choice-set-independent MNP model (MNP_{csi}). For this MNP model, the assumption of utility maximization results in an expression for π_{km} that involves an (*M*-1)-dimensional integral:

$$(18) \qquad \pi_{km} = P(u_{kn} - u_{km} \leq 0 \ \forall \ n \neq m \in \Delta_k) = P(\tilde{u}_{km} \leq 0) = \int_{-\infty}^{0} d_{km}(t)\, dt$$

where Δ_k is the set of profiles in choice set k and $d_{km}(.)$ is the density of $\tilde{u}_{km}$. This specification accounts for local contrast effects only, since it allows utilities within choice sets to be correlated.

Table 5: MNP_{csi} Estimation Results

Attribute (level)	β Est.	β S.e.	σ Est.	σ S.e.
1 Brand (1)	-0.106	.203	0.717	.386
2 Brand (2)	-0.179	.222	0.107	.665
3 Capacity (1)	-1.166*	.123	0.585*	.270
4 Capacity (2)	0.587*	.090	-0.001	.224
5 Price (1)	0.326	.497	-0.084	.599
6 Price (2)	0.378	.269	0.482	.473
7 Filter (1)	0.354*	.098	0.298	.372
8 Thermos (1)	0.269	.152	0.173	.259
Statistics				
Ln-likelihood	-1279.100			
AIC	2590.201			
BIC	2674.997			
Pseudo R^2	0.213			

*: $p < 0.05$.

In Table 5 the coffee-maker data results are listed for the structural parameters β and the covariance parameters σ, as well as the fit-statistics. It shows that the results of the MNP_{csi} model are a somewhat disappointing. Although it produces a better log-likelihood than the MNL model, due to the high standard errors only a few parameters are significantly different from zero with a p-value of 5%. The results indicate that the capacity and presence of a special filter are the only relevant attributes. Furthermore, the one significant covariance parameter (belonging to the first capacity level) is responsible for the increased fit of the MNP_{csi} model with respect to the MNL model, but no clear conclusions can be drawn from these results. In addition, the AIC and BIC statistics indicate that the LCMNL has better fit than the MNP_{csi} model. The estimates of Table 5 were use to predict the hold-out sets. The predicted log-likelihood is equal to -784.677 (Pseudo R^2 = 0.035), which is worse from the LCMNL model and even worse than those of the MNL model.

The results of the MNP_{csi} model indicate that allowing for heterogeneity and correlation of utilities within choice sets may help to improve model fit in terms of the log-likelihood value. However, this application also showed that the discrete (LCMNL) representation of heterogeneity seems to do better than the continuous (MNP) one. In the next subsection the MNP model that in addition allows for correlations between choice sets is developed.

Multinomial Probit with dependent choice sets

The MNP specification in the previous subsection only allowed for correlations within choice sets and not between. In this section we assume that utilities of the same individual are not independent over choice sets, but rather that utilities of alternatives in different choice sets are correlated. In this case a total probability has to be obtained for the complete choice observation of an individual. A simple example illustrates this. Assume we have two choice sets with each three alternatives (so, H=5). For each individual we observe two choices, one from each set. Consider an individual j choosing the second alternative from the first set and the base alternative from the second choice set. The resulting joint probability for this example is equal to (b represents the base alternative):

(19) $$\pi_{2b} = P(\, u_{j12} > u_{j11} \;, u_{j12} > u_{j1b} \;, u_{j2b} > u_{j21} \;, u_{j2b} > u_{j22}\,).$$

This probability can be expressed involving a four-dimensional integral. In the general case, a K-vector of choices is observed for each individual, and we have to consider M^K arrays containing the multiple choices from different choice sets. Each array corresponds to a joint probability, involving an (H-1)-dimensional integral that describes the probability of observing the array of choices from all choice sets (cf. Hausman and Wise 1978; Papatla 1996). In this case, the form of the probabilities for the MNP model becomes somewhat complicated, and we omit the formal presentation of these probabilities since the notation provides no additional insight.

The log-likelihood for this MNP approach is equal to:

(20) $$L_{MNP} = J \sum_{l=1}^{M^K} p_l \ln(\pi_l),$$

where l indexes the K-dimensional choice arrays, p_l denotes the observed fractions of the choice arrays, and π_l denotes the choice probabilities expressed as functions of the model parameters. This specification accounts for both the local and background contrast effect, because the choice probabilities, as in (19), depend on all profiles in the design or, alternatively, with heterogeneity of the parameters across choice sets. This is not the case with models that treat the choice sets as independent, such as the MNL model, LCMNL model and the MNP model of the previous subsection.

Table 6 lists the parameter estimates and fit statistics. After estimation, all eigenvalues of the final Hessian were positive, indicating that the model is identified (Bekker, Merckens and Wansbeek 1994).

Table 6 shows the same pattern of β-estimates as in the MNL model. However, the fit of the MNP model is much better than that of the MNL model. This is caused by the estimated covariance parameters. The log-likelihood of the MNP model is somewhat lower than that of the LCMNL model, and consequently also the Pseudo R^2 is lower. The AIC statistic would favor the LCMNL model over the

MNP model, but the BIC statistic, which penalties the likelihood more severely, indicates the MNP model as best. Table 6 shows that most attribute levels are potentially responsible for correlations between and within choice sets. Note, however, that because of the effects-type coding some effects may cancel out (see Haaijer (1999), chapter 5.2.4 for a discussion), so one has to investigate not only the estimated parameters but in addition the estimated covariance matrix (16).

Table 6: MNP Estimation Results

Attribute (level)	β		σ	
	Est.	*S.e.*	*Est.*	*S.e.*
1 Brand (1)	-0.029	.101	0.417*	.096
2 Brand (2)	-0.240*	.078	-0.387*	.099
3 Capacity (1)	-1.075*	.092	0.850*	.094
4 Capacity (2)	0.565*	.060	-0.348*	.083
5 Price (1)	0.432*	.116	-0.562*	.139
6 Price (2)	0.244*	.082	-0.145	.100
7 Filter (1)	0.355*	.038	0.023	.058
8 Thermos (1)	0.393*	.054	-0.206*	.071
Statistics				
Ln-likelihood	-1086.622			
AIC	2205.245			
BIC	2256.770			
Pseudo R^2	0.332			

*: $p < 0.05$.

Interestingly, the covariance matrix of the MNP model reveals alternatives with near zero covariances with all other alternatives. This indicates that these are (almost) independent of the other alternatives. After constructing the Ω matrix the σ-estimates reveal what attribute (levels) are responsible for correlations within and between choice sets.

The estimates of the MNP model were again used to predict the holdout sets. For the MNP model the predicted log-likelihood is equal to -679.075 (Pseudo R^2 =0.165). This shows that it is very important to account for both kind of correlations, or context effects, which not only results in an improved model fit, with relatively few covariance parameters, but also in an improved holdout predictive fit, which is better than that of the MNL, MNP_{csi}, and LCMNL models.

11.5 Discussion and Conclusion

In the above sections we showed several models that can be used to analyze conjoint choice experiments. We discussed the standard MNL model, the Latent Class

MNL model and two versions of an MNP model with a specific covariance structure. The performance of the models was illustrated with an application. It is interesting to compare the results of the various models, although we realize that this comparison is based on only this one application, so some care must be taken with respect to drawing conclusions.

When we compare the estimated log-likelihood values, the LCMNL model gives the best result, followed by the MNP model. At some distance the MNP_{csi} model and MNL model follow. However, since the number of parameters in the LCMNL model is much larger than that in the MNP model (35 and 16 respectively), the BIC-statistic, that compensates for the number of parameters and observations in the model, ranks the MNP model as best. The AIC-statistic, that only puts a penalty on the number of parameters, still lists the LCMNL model as best. If we compare the predictive power of the models the MNP model comes out as best, followed by the LCMNL, MNL and MNP_{csi} models. Based on these results one could conclude that the using the MNL model for (conjoint) choice experiments may not be a good choice. It does not account for heterogeneity and correlated choice alternatives within and between choice sets. This results in an inferior model fit and predictive fit compared to models that do account for these elements. On the positive side, however, we saw that the estimates for the structural parameters do not differ much between the models. This is in line with findings by Börsch-Supan et al. (1990), who also found that differences in model fit, with respect to the log-likelihood value, is often caused by the error structure while the structural parameters are relatively left unaffected, although a misspecified covariance matrix not only affects the standard errors of the covariance parameters but also of the structural parameters. This is exactly what may have happened in the MNP_{csi} model. So, in terms of finding the „right" structural parameters the MNL model may do a reasonable job, but when these results are used to predict the performance of new alternatives, or holdout choice sets, the MNL model falls short to the LCMNL and MNP model. Of course, further research should be done to test the generalizability of these findings. Haaijer et al. (1998) tested three data sets, including the one in this chapter, and found in all three cases that the MNP model outperforms the Independent Probit model (which has similar characteristics as the MNL model) on estimation fit and holdout predictive power. Haaijer, Kamakura and Wedel (2000) support these findings for two other data sets.

Several authors have compared continuous and discrete specifications of heterogeneity (e.g., Lenk, DeSarbo, Green and Young 1996; Vriens, Wedel and Wilms 1996; Allenby, Arora and Ginter 1998; Allenby and Rossi 1999). These comparisons were made on scanner panel data rather than on conjoint choice data. Nevertheless, the conclusion from these studies is that for predictive purposes continuous (MNP) specifications may be preferable over discrete (LCMNL) specifications. From a substantive angle, the MNP-type of specification, particularly when applied in conjunction with the Gibbs sampler, that allows for individual level parameters to be estimated, seems preferable in direct marketing applications, where such individual level estimates are of great use in targeting individuals. However, advantages of the discrete model specification accrue in situations where managers are interested in targeting market segments (see Wedel and Kamakura 1997 for an extensive discussion).

Several other issues need to be further investigated. The findings of this chapter that local and background context play a role in respondents' choices should be studied more closely. The importance of these effects in relation to the number of attributes, choice sets, levels of attributes and alternatives could be given additional attention. The attribute level effect found in the literature could be particularly related to context effects. Furthermore, the involvement of respondents and their knowledge on the product category, as well as the product category itself, could also influence the importance of these context effects. The influence of all these factors on the choice, and hence ultimately on the parameters of interest, should be minimized or at least be accounted for. Research should be done on the optimal design for conjoint choice experiments analyzed with Probit models, since optimal designs are not yet available. The results obtained from such an optimal conjoint experiments should lead to managerial more insightful and precise information on the product in question. In addition, one needs research on many product categories, designs, etcetera to investigate how generalizable findings are.

Furthermore, the performance of the Simulated Maximum Likelihood method explored in this chapter as optimization methods for the MNP models should be compared with Bayesian estimation using the Gibbs sampler. An interesting avenue for further research is in combining the Latent Class and MNP approaches, thus having the advantages of the predictive performance of the continuous and of the managerial appeal of segments of the discrete heterogeneity representation. It would be interesting to test the performance of such a Latent Class MNP model. In this chapter we already saw that the MNP model leads to better prediction results as compared to the LCMNL model. In addition accounting for different segments in an MNP context could further improve predictive performance and enhance managerial appeal, although the number of parameters to estimate may become a limiting factor, with respect to estimation time as well as their identification.

In any case, we may conclude that the MNL (or IP) model is no longer the preferred choice for analyzing conjoint choice experiments.

11.6 References

Addelman, S. (1962), Orthogonal Main-Effects Plans for Asymmetrical Factorial Experiments, *Technometrics*, 4, 21-46.

Akaike, H. (1973), Information Theory and an Extension of the Maximum Likelihood Principle, in: Petrov, B.N. and Csáki, F., eds., *2nd International Symposium on Information Theory*, Akadémiai Kiadó, Budapest, 267-281.

Allenby, G. M., Arora, N. and Ginter, J. L. (1995), Incorporating Prior Knowledge into the Analysis of Conjoint Studies, *Journal of Marketing Research*, 32, 152-162.

Allenby, G. M., Arora, N. and Ginter, J. L. (1998), On the Heterogeneity of Demand, *Journal of Marketing Research*, 35, 384-389.

Allenby, G. M. and Ginter, J. L. (1995), Using Extremes to Design Products and Segment Markets, *Journal of Marketing Research*, 32, 392-403.

Allenby, G. M. and Rossi, P. E. (1999), Marketing Models of Consumer Heterogeneity, *Journal of Econometrics*, 89, 57-78.

Arora, N., Allenby, G. M. and Ginter, J. L. (1998), A Hierarchical Bayes Model of Primary and Secondary Demand, *Marketing Science*, 17, 29-44.

Bekker, P. A., Merckens, A. and Wansbeek, T. J. (1994), *Identification, Equivalent Models, and Computer Algebra*, San Diego.

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

Börsch-Supan, A., Hajivassiliou, V. A., Kotlikoff, L. J. and Morris J. N. (1990), Health, Children, and Elderly Living Arrangements: A Multiperiod, Multinomial Probit Model with Unobserved Heterogeneity and Autocorrelated Errors, NBER working Paper 3343.

Bunch, D. S. (1991), Estimability in the Multinomial Probit Model, *Transportation Research B*, 25, 1-12.

Bunch, D. S. and Kitamura, R. (1989), Multinomial Probit Model Estimation Revisited: Testing Estimable Model Specifications, Maximum Likelihood Algorithms, and Probit Integral Approximations for Trinomial Models of Household Car Ownership, working paper, University of California at Davis.

Carroll, J. D. (1972), Individual Differences and Multidimensional Scaling, in: Shepard, R. N., Romney, A. K. and Nerlove, S. B., eds., *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences*, New York.

Carroll, J. D. and Green, P. E. (1995), Psychometric Methods in Marketing Research: Part 1, Conjoint Analysis, *Journal of Marketing Research*, 32, 385-391.

Cattin, P. and Wittink, D. R. (1982), Commercial Use of Conjoint Analysis: A Survey, *Journal of Marketing*, 46, 44-53.

Chintagunta, P. K. (1992), Estimating A Multinomial Probit Model Of Brand Choice Using The Method Of Simulated Moments, *Marketing Science*, 11, 386-407.

Chrzan, K. (1994), Three Kinds of Order Effects in Choice-Based Conjoint Analysis, *Marketing Letters*, 5, 165-172.

Cohen, S. H. (1997), Perfect Union: CBCA Marries the Best of Conjoint and Discrete Choice Models, *Marketing Research*, 9, 12-17.

Daganzo, C. F. (1979), *Multinomial Probit, The Theory and Its Applications to Demand Forecasting*, New York.

Dansie, B. R. (1985), Parameter Estimability in the Multinomial Probit Model, *Transportation Research B*, 19, 526-528.

Dellaert, B. G. C., Borgers, A. W. J. and Timmermans, H. J. P. (1995), A Day in the City, Using Conjoint Choice Experiments to Model Tourists Choice of Activity Packages, *Tourism Management*, 16, 347-353.

Dellaert, B. G. C., Borgers, A. W. J. and Timmermans, H. J. P. (1996), Conjoint Choice Models of Joint Participation and Activity Choice, *International Journal of Research in Marketing*, 13, 251-264.

Dellaert, B. G. C., Borgers, A. W. J. and Timmermans, H. J. P. (1997), Conjoint Models of Tourist Portfolio Choice: Theory and Illustration, *Leisure Science*, 19, 31-58.

DeSarbo, W. S. and Green, P. E. (1984), Concepts, Theory, and Techniques, Choice-Constrained Conjoint Analysis, *Decision Science*, 15, 291-323.

DeSarbo, W. S., Ramaswamy, V. and Cohen, S. H. (1995), Market Segmentation with Choice-Based Conjoint Analysis, *Marketing Letters*, 6, 137-148.

Elrod, T. and Keane, M. P. (1995), A Factor-Analytic Probit Model for Representing the Market Structures in Panel Data, *Journal of Marketing Research*, 32, 1-16.

Elrod, T., Louviere, J. J. and Davey, K. S. (1992), An Empirical Comparison of Rating-Based and Choice-Based Conjoint Models, *Journal of Marketing Research*, 29, 368-377.

Green, P. E. (1984), Hybrid Models for Conjoint Analysis: An Expository Review, *Journal of Marketing Research*, 21, 155-169.

Green, P. E., Goldberg, S. M. and Montemayor, M. (1981), A Hybrid Utility Estimation Model for Conjoint Analysis, *Journal of Marketing*, 45, 33-41.

Green, P. E. and Rao, V. R. (1971), Conjoint Measurement for Quantifying Judgmental Data, *Journal of Marketing Research*, 8, 355-363.

Green, P. E. and Srinivasan, V. (1978), Conjoint Analysis in Consumer Research: Issues and Outlook, *Journal of Consumer Research*, 5, 103-123.

Green, P. E. and Srinivasan, V. (1990), Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice, *Journal of Marketing*, 54, 3-19.

Haaijer, M. E. (1999), Modeling Conjoint Choice Experiments with the Probit Model, University of Groningen.

Haaijer, M. E., Kamakura, W. A. and Wedel, M. (2000), Response Latencies in the Analysis of conjoint Choice Experiments, Journal of Marketing Research, 37, 376-382.

Haaijer, M. E., Kamakura, W. A. and Wedel, M. (2001), The No-Choice Alternative in Conjoint Choice Experiments, International Journal of Market Research, 43, 93-106.

Haaijer, M. E., Wedel, M., Vriens, M. and Wansbeek, T. J. (1998), Utility Covariances and Context Effects in Conjoint MNP Models, *Marketing Science*, 17, 236-252.

Hajivassiliou, V. A. (1993), Simulation Estimation Methods for Limited Dependent Variable Models, in: Maddala, G. S., Rao, C. R. and Vinod, H.D., eds., *Handbook of Statistics*, 11, 519-543.

Hausman, J. A. and Wise, D. A. (1978), A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences, *Econometrica*, 46, 403-426.

Huber, J., Payne, J. W. and Puto, C. (1982), Adding Asymmetrically Dominant Alternatives: Violations of Regularity and the Similarity Hypotheses, *Journal of Consumer Research*, 9, 90-98.

Huber, J. and Puto, C. (1983), Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects, *Journal of Consumer Research*, 10, 31-44.

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

Johnson, R. M. (1974), Trade-off Analysis of Consumer Values, *Journal of Marketing Research*, 11, 1221-1227.

Johnson R. M. (1985), Adaptive Conjoint Analysis, in: *Proceedings of the Sawtooth Software Conference on Perceptual Mapping, Conjoint Analysis and Computer Interviewing*, Ketchum, ID: Sawtooth Software, Inc., 253-265.

Johnson, R. M. and Orme, B. K. (1996), How Many Questions Should You Ask In Choice-Based Conjoint Studies?, Sawtooth Software Technical Paper.

Kamakura, W. A. (1989), The Estimation of Multinomial Probit Models: A New Calibration Algorithm, *Transportation Science*, 23, 253-265.

Kamakura, W. A., Wedel, M. and Agrawal, J. (1994), Concomitant Variable Latent Class Models for Conjoint Analysis, *International Journal of Research in Marketing*, 11, 451-464.

Keane, M. P. (1992), A Note on Identification in the Multinomial Probit Model, *Journal of Business & Economic Statistics*, 10, 193-200.

Kruskal, J. B. (1965), Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data, *Journal of the Royal Statistical Society*, Series B, 251-263.

Kuhfeld, W. F., Tobias, R. D. and Garratt, M. (1994), Efficient Experimental Design with Marketing Research Applications, *Journal of Marketing Research*, 31, 545-557.

Lenk, P. J., DeSarbo, W. S., Green, P. E. and Young, M. R. (1996), Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs, *Marketing Science*, 15, 173-191.

Louviere, J. J. (1988), Conjoint Analysis Modeling of Stated Preferences. A Review of Theory, Methods, Recent Developments and External Validity, *Journal of Transport Economics and Policy*, 10, 93-119.

Louviere, J. J. and Woodworth, G. (1983), Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data, *Journal of Marketing Research*, 20, 350-367.

Louviere, J. J. and Timmermans, H. J. P. (1990), A Review of Recent Advances in Decompositional Preference and Choice Models, *Tijdschrift voor Economische en Sociale Geografie*, 81, 214-225.

Luce, R. D. and Tukey, J. W. (1964), Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement, *Journal of Mathematical Psychology*, 1, 1-27.

Madansky, A. (1980), On Conjoint Analysis and Quantal Choice Models, *Journal of Business*, 53, S37-S44.

Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.

McFadden, D. (1976), Quantal Choice Analysis: A Survey, *Annals of Economic and Social Measurement*, 5, 363-390.

McFadden, D. (1986), The Choice Theory Approach to Market Research, *Marketing Science*, 5, 275-297.

Moore, W. L., Gray-Lee, J. and Louviere, J. J. (1998), A Cross-Validity Comparison of Conjoint Analysis and Choice Models at Different Levels of Aggregation, *Marketing Letters*, 9, 195-207.

Nowlis, S. M. and Simonson, I. (1997), Attribute-Task Compatibility as a Determinant of Consumer Preference Reversals, *Journal of Marketing Research*, 34, 205-218.

Oliphant, K., Eagle, T. C., Louviere, J. J. and Anderson, D. (1992), Cross-Task Comparison of Rating-Based and Choice-Based Conjoint, Proceedings of the Sawtooth Software Conference 1992.

Oppewal, H., Louviere, J. J. and Timmermans, H. J. P. (1994), Modeling Hierarchical Conjoint Processes with Integrated Choice Experiments, *Journal of Marketing Research*, 31, 92-105.

Oppewal, H. and Timmermans, H. J. P. (1993), Conjuncte Keuze-Experimenten: Achtergronden, Theorie, Toepassingen en Ontwikkelingen, *NVVM Jaarboek*, 33-58. (In Dutch).

Papatla, P. (1996), A Multiplicative Fixed-effects Model of Consumer Choice, *Marketing Science*, 15, 243-261.

Pekelman, D. and Sen, S. (1974), Mathematical Programming Models for the Determination of Attribute Weights, *Management Science*, 20, 1217-1229.

Pudney, S. (1989), *Modeling Individual Choice: The Econometrics of Corners, Kinks and Holes*, Basil Blackwell Inc., Oxford.

Rossi, P. E., McCulloch, R. E. and Allenby, G. M. (1996), The Value of Purchase History Data in Target Marketing, *Marketing Science*, 15, 321-340.

Sándor, Z. and Wedel, M. (1999), Robust Optimal Designs for Conjoint Choice Experiments, *SOM-Working Paper*, University of Groningen, Netherlands.

Sawtooth Software Inc. (1995), The CBC System for Choice-Based Conjoint Analysis, Sawtooth Software Technical Paper.

Schwarz, G. (1978), Estimating the Dimension of a Model, *Annals of Statistics*, 6, 461-464.

Simonson, I. (1989), Choice Based on Reasons: The Case of Attraction and Substitution Effects, *Journal of Consumer Research*, 16, 158-174.

Simonson, I. and Tversky, A. (1992), Choice in Context: Tradeoff Contrasts and Extremeness Aversion, *Journal of Marketing Research*, 29, 281-295.

Srinivasan, V. and Shocker, A. D. (1973a), Linear Programming Techniques for Multidimensional Analysis of Preferences, *Psychometrika*, 38, 337-369.

Srinivasan, V. and Shocker, A. D. (1973b), Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgements, *Psychometrika*, 38, 473-493.

Steenkamp, J. E. B. M. (1985), De Constructie van Profielensets voor het Schatten van Hoofdeffecten en Interacties bij Conjunct Meten, *NVVM Jaarboek*, 125-155. (In Dutch).

Struhl, S. (1994), Discrete Choice Modelling: Understanding a Better Conjoint Than Conjoint, *Quirk's Marketing Research Review*, 36-39.

Swait, J. and Louviere, J. J. (1993), The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models, *Journal of Marketing Research*, 30, 305-314.

Timmermans, H. J. P. and Van Noortwijk, L. (1995), Context Dependencies in Housing Choice Behavior, *Environment and Planning A*, 27, 181-192.

Tversky, A. (1972), Elimination by Aspects: A Theory of Choice, *Psychological Review*, 79, 281-299.

Vriens, M. (1994), Solving Marketing Problems With Conjoint Analysis, *Journal of Marketing Management*, 10, 37-55.

Vriens, M. (1995), *Conjoint Analysis in Marketing, Developments in Stimulus Representation and Segmentation Methods*, Thesis, University of Groningen.

Vriens, M., Oppewal, H. and Wedel, M. (1998), Ratings-Based versus Choice-Based Latent Class Conjoint Models - An Empirical Comparison, *Journal of the Market Research Society*, 40, 237-248.

Vriens, M., Wedel, M. and Wilms, T. (1996), Segmentation Methods for Metric Conjoint Analysis: A Monte Carlo Comparison, *Journal of Marketing Research*, 33, 73-85.

Wedel, M. and Kamakura, W.A. (1997), *Market Segmentation: Conceptual and Methodological Foundations*, Dordrecht, Kluwer.

Wedel, M., Kamakura, W., Arora, N., Bemmaor, A., Chiang, J., Elrod, T., Johnson, R., Lenk, P., Neslin, S. and Poulsen, C.S. (1999), Heterogeneity and Bayesian Methods in Choice Modeling, working paper, University of Groningen.

Wedel, M., Vriens, M., Bijmolt, T. H. A., Krijnen, W. and Leeflang, P. S. H. (1998), Assessing the Effects of Abstract Attributes and Brand Familiarity in Conjoint Choice Experiments, *International Journal of Research in Marketing*, 15, 71-78.

Wittink, D. R. and Cattin, P. (1989), Commercial Use of Conjoint Analysis: An Update, *Journal of Marketing*, 53, 91-96.

Wittink, D. R., Huber, J., Fiedler, J. A. and Miller, R. L. (1991), The Magnitude of and an Explanation/Solution for the Number of Levels Effect in Conjoint Analysis, working paper, Cornell University.

Wittink, D. R., Vriens, M. and Burhenne, W. (1994), Commercial Use of Conjoint Analysis in Europe: Results and Critical Reflections, *International Journal of Research in Marketing*, 11, 41-52.