Utility Covariances and Context Effects in Conjoint MNP Models

Rinus Haaijer • Michel Wedel • Marco Vriens • Tom Wansbeek
University of Groningen, Faculty of Economics, Department of Marketing and Marketing Research, PO Box 800, 9700 AV Groningen, The Netherlands
m.e.haaijer@eco.rug.nl
University of Groningen, Groningen, The Netherlands
Research International U.S.A., San Francisco, California
University of Groningen, Groningen, The Netherlands

Abstract

Experimental conjoint choice analysis is among the most frequently used methods for measuring and analyzing consumer preferences. The data from such experiments have been typically analyzed with the Multinomial Logit (MNL) model. However, there are several problems associated with the standard MNL model because it is based on the assumption that the error terms of the underlying random utilities are independent across alternatives, choice sets, and subjects. The Multinomial Probit model (MNP) is well known to alleviate this assumption of independence of the error terms. Accounting for covariances in utilities in modeling choice experiments with the MNP is important because variation of the coefficients in the choice model may occur due to context effects. Previous research has shown that subjects' utilities for alternatives depend on the choice context, that is, the particular set of alternatives evaluated. Simonson and Tversky's tradeoff contrast principle describes the effect of the choice context on attribute importance and patterns of choice. They distinguish *local contrast effects*, which are caused by the alternatives in the offered set only, and *background contrast effects*, which are due to the influence of alternatives previously considered in choice experiments. These effects are hypothesized to cause correlations in the utilities of alternatives within and across choice sets, respectively.

The purpose of this study is to develop an MNP model for conjoint choice experiments. This model is important for a more detailed study of choice patterns in those experiments. In developing the MNP model for conjoint choice experiments, several hurdles need to be taken related to the identification of the model and to the prediction of holdout profiles. To overcome those problems, we propose a random coefficients (RC) model that assumes a multivariate normal distribution of the regression coefficients with a rank one factor structure on the covariance matrix of these regression coefficients. The parameters in this covariance matrix can be used to identify which attributes and levels of attributes are potentional sources of dependencies between the alternatives and choice sets in a conjoint choice experiment. We present several versions of this model. Moreover, for each of these models we allow utilities to be either correlated or independent across choice sets. The Independent Probit (IP) model is used as a benchmark. Given the dimensionality of the integrations involved in computing the choice probabilities, the models are estimated with simulated likelihood, where simulations are used to approximate the integrals involved in the choice probabilities.

We apply and compare the models in two conjoint choice experiments. In both applications, the random coefficients MNP model that allows choices in different choice sets to be correlated (RC) displays superior fit and predictive validity compared with all other models. We hypothesize that the difference in fit occurs because the RC model accommodates correlations among choice sets that are caused by background contrast effects, whereas the model that treats choice sets as independent (iRC) accounts for local contrast effects only. The iRC model shows superior model fit compared with the IP model, but its predictions are worse than those of the IP model. We find differences in the importance of local and background contrast effects for choice sets containing different numbers of alternatives: The background contrast effect may be stronger for smaller choice sets, whereas the local contrast effect may be stronger for bigger choice sets.

We illustrate the differences in simulated market shares that are obtained from the RC, iRC, and IP models in three hypothetical situations: product modification, product line extension, and the introduction of a me-too brand. In all of those situations, substantially different market shares are predicted by the three models, which illustrates the extent to which erroneous predictions may be obtained from the misspecified iRC and IP models.
(*Conjoint Choice Experiments*; *Multinomial Probit*; *Random Taste Variation*; *Random Utility*)

Marketing Science/Vol. 17, No. 3, 1998
pp. 236–252
0732-2399/98/1703/0236$05.00

1. Introduction

Experimental conjoint choice analysis is among the most frequently used methods for measuring and analyzing consumer preferences. In particular, it has been applied with great success in the fields of new product design (Louviere and Woodworth 1983), product optimization (Green and Krieger 1993), and the evaluation of competitive sets (Mahajan, Green, and Goldberg 1982; Louviere and Woodworth 1983). 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. The 0/1 choice data arising from such conjoint choice experiments have been typically analyzed with the Multinomial Logit (MNL) model (e.g., Louviere and Woodworth 1983; Elrod, Louviere, and Davey 1992). In their *Journal of Marketing Research* editorial on conjoint analysis, Carroll and Green (1995) point to some of the advantages of experimental choice analysis 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 (cf. Louviere 1988; Cohen 1997).

There are several problems associated with the standard MNL model used to analyze the choice experiments, however. First, it does not deal with consumer heterogeneity (cf. Allenby, Arora, and Ginter 1995; Carroll and Green 1995; Chintagunta and Honoré 1996; Cohen 1997; Keane 1997). In addition, problems arise because it is derived from random utility maximization, based on the assumption that the error terms are independent across alternatives, choice sets, and subjects (e.g., McFadden 1976; Hausman and Wise 1978; Currim 1982; Kamakura and Srivastava 1984). This leads to the property of Independence of Irrelevant Alternatives (IIA), where little is known about its validity in conjoint choice experiments (Carroll and Green 1995).

It is well known that the Multinomial Probit model (MNP) relaxes the assumption of independence of the error terms in random utility models (e.g., Daganzo 1979; Kamakura 1989; Chintagunta 1992) and thereby alleviates IIA. In the conjoint choice context, the MNP model offers the major advantage of allowing correlations among the random utilities of alternatives within choice sets and among the repeated choices that consumers make from the multiple-choice sets. We will show that the MNP model accounts for heterogeneity in the coefficients occurring due to context effects in conjoint choice experiments.

Heterogeneity of the coefficients of the choice model across the population of consumers may cause utilities to be correlated (cf. Daganzo 1979). Estimating models at the individual level, or including consumer characteristics in the model, cannot be recommended as general solutions to that problem (Vriens, Wedel, and Wilms 1996). The finite mixture approach to modeling heterogeneity in choice experiments (DeSarbo, Ramaswamy, and Cohen 1995) has been criticized by some authors (Allenby and Lenk 1994; Allenby and Ginter 1995), who argue that tastes follow a continuous distribution over the population rather than a discrete one, or that the assumption of within-segment homogeneity in finite mixture models may result in a loss of predictive performance because predictions are restricted to be a convex combination of segment-level parameters (Lenk, DeSarbo, Green, and Young 1996). The issue of the appropriateness of discrete or continuous parametric representations is, however, largely empirical. In this paper we opt for a continuous representation. The MNP model allows heterogeneity or random taste variation to be included by assuming a continuous distribution of the attribute-level coefficients across the population (e.g., Hausman and Wise 1978; Papatla 1996). Chintagunta and Honoré (1996) concluded from their study that allowing for such heterogeneity constitutes a major improvement in both the fit and the substantive implications derived from discrete choice models. However, with sufficient support points, a discrete approach should approximate a continuous distribution quite well.

A problem that seems to have received limited attention in the experimental choice modeling literature is that of heterogeneity in the coefficients of the choice model due to context effects. Empirical evidence has accumulated indicating that subjects' utilities for alternatives depend on the choice context (e.g., Huber, Payne, and Puto 1982; Simonson and Tversky 1992; Nowlis and Simonson 1997), where "context" is defined as the particular set of alternatives evaluated. These studies have shown that the tradeoff of attributes, reflected in the importance subjects attach to them, vary according to the choice context. Since the design of experimental 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) proposed the tradeoff contrast principle to describe the effect of the choice context on attribute importance and patterns of choice. They 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. Local contrast effects may occur in a conjoint choice experiment 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). Background contrast effects may occur in conjoint choice experiments if the attribute importance of profiles in a particular choice set is influenced by tradeoffs among profiles in previous choice sets. This leads to covariance among the random utilities of alternatives in different choice sets and violates the assumption of independence of choices among alternatives in different sets in MNL models. More specific context effects hypotheses have been formulated, such as substitution, dominance, attraction, and compromise effects (Tversky 1972, Huber, Payne, and Puto 1982; Huber and Puto 1983; Simonson 1989; Simonson and Tversky 1992). A discussion of these effects is, however, beyond the scope of the present paper.

In § 2 we develop the MNP model for conjoint choice experiments. It accounts for covariances among the alternatives and choice sets potentially caused by heterogeneity and context effects. Several hurdles need to be taken in developing the MNP model. These hurdles are related to the dimensionality of the integrations involved in computing the choice probabilities, to the identification of the model, and to the prediction of holdout profiles. To overcome those problems, we propose a specific form of the MNP model that imposes a factor structure on the covariance matrix and is estimated with simulation. We present several restricted versions of the MNP model that allow us to explore the correlational structure of random utilities within and across choice sets. In § 3 we present two applications to conjoint choice experiments on cars and coffee makers. We compare several versions of the model and empirically demonstrate the advantages of accounting for the covariance among utilities: substantially improved model fit and predictive accuracy. We also illustrate differences between the various MNP models in performing market share simulations for three strategic scenarios that apply to the coffee-makers application: product modification, product line extension, and the introduction of me-too brands. In § 4 we present discussion and conclusions.

2. Conjoint Choice MNP Models

In this section we present the various MNP models that we will estimate later. In § 2.1 we specify the general MNP model for conjoint choice experiments. In § 2.2 we develop the specification for the covariance matrix and propose several alternative model formulations.

2.1. Multinomial Probit with Multiple-Choice Sets

In this section we focus for convenience on the particular choice set design used in the empirical applications that follow. Assume there are J respondents and H different alternatives (profiles), which are divided into K smaller sets with M alternatives each. One profile, labeled (arbitrarily) 1, is the base alternative that is common to all sets and that scales the utility levels between choice sets. The other profiles are unique to their particular choice set, so that $H = K(M - 1) + 1$. It is possible to accommodate a more general choice design, for instance with other profiles besides the base alternative that appear in more than one choice set or choice sets with different sizes, but we restrict ourselves to the situation described above for simplicity. The utilities of the alternatives for individual j are contained in the latent unobservable vector u_j, which satisfies

$$u_j = X\beta_j + e_j, \tag{1}$$

where X is a $(H \times S)$ matrix containing the attributes of the alternatives, β_j is a $(S \times 1)$ vector of random weights, and e_j is the vector containing the random component of the utilities, assumed to be distributed as

$$e_j \sim N_H(0, \Sigma_e), \tag{2}$$

independent between individuals; Σ_e is a $(H \times H)$ positive definite covariance matrix. We allow for random variation in the attribute level coefficients by specifying β_j in Equation (1) as (cf., e.g., Hausman and Wise 1978; Daganzo 1979; Ben-Akiva and Lerman 1985)

$$\beta_j = \beta + \psi_j, \tag{3}$$

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

$$u_j \sim N_H(X\beta, \Omega), \tag{4}$$

with

$$\Omega = \Sigma_e + X\Sigma_\psi X'. \tag{5}$$

We consider two ways to deal with the conjoint data structure. First, we take an individual's utilities to be independent between the choice sets. We then have JK independent observations, and the log-likelihood is a straightforward generalization of the standard likelihood of choice models, where a summation over choice sets is introduced. Letting p_{km} denote the fraction of individuals choosing alternative m in set k, the log-likelihood is

$$L_1 = J \sum_{k=1}^{K} \sum_{m=1}^{M} p_{km} \ln(\pi_{km}), \tag{6}$$

where π_{km} is the probability that alternative m is chosen in set k. Note 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. For the MNP model, the assumption of utility maximization results in an expression for π_{km} that involves an $(M - 1)$-dimensional integral:

$$\pi_{km} = P(u_{kn} - u_{km} \leq 0 \ \forall \ n \neq m \in \Delta_k)$$
$$= P(\tilde{u}_{km} \leq 0) = \int_{-\infty}^{0} d_{km}(t)dt, \tag{7}$$

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

Second, one may instead 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.[1] In this case, the form of the likelihood is more complicated. The consequences are most easily illustrated by a simple example. Let $K = 2$ and $M = 3$ $(H = 5)$, hence the two choice sets have indexes {1, 2, 3} and {1, 4, 5}, respectively. For each individual we observe two choices, one from each set. Consider an individual j choosing "2" from set 1 and "4" from set 2. The resulting joint probability for this example is equal to

$$\pi_{24} = P(u_{j2} > u_{j1}, u_{j2} > u_{j3}, u_{j4} > u_{j1}, u_{j4} > u_{j5}). \tag{8}$$

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). We omit a formal presentation of this probability because the notation is extremely burdensome without providing additional insight. The log-likelihood for this approach is equal to

$$L_2 = J \sum_{l=1}^{M^K} p_l \ln(\pi_l), \tag{9}$$

where l indexes the K-dimensional choice arrays, p_l denotes fractions of the choice arrays, and π_l denotes the choice probabilities expressed as functions of the model parameters. In this model we account in addition for the background contrast effect because the choice probabilities, as in (8), 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 and the MNP model of (6) and (7).

[1]We are grateful to an anonymous reviewer of an earlier version of our paper for pointing this out to us.

Estimates for the parameters are obtained by maximization of the likelihood in (6) or (9) over β and the parameters in the covariance matrix. However, when the dimension of the integral is greater than three, the probabilities π_{km} in (6) or π_l in (9) cannot be evaluated numerically (e.g., McFadden 1976; Maddala 1983; Kamakura 1989; Keane 1992). The $(H - 1)$-dimensional integrals (*one* for each respondent) involved in the likelihood L_2 (9) and the $(M - 1)$-dimensional integrals (K for each respondent) in L_1 (6) are approximated using simulation. We use the SRC simulator to apply the Simulated Maximum Likelihood (SML) method. Details are given in the appendix. Note that in conjoint choice experiments the number of respondents, J, is often much lower than the number of possible different choice arrays, M^K, so in practical applications only a maximum of J probabilities have to be simulated in the maximization of (9).

2.2. Random Intercept and Random Coefficient Models

In the general MNP model provided by Equations (1) and (2), identification requires certain restrictions on the parameters of the covariance matrix (e.g., Bunch 1991; Bunch and Kitamura 1991; Keane 1992). When choice sets contain M alternatives, $M + 1$ covariance parameters should be fixed, or alternatively, one variance should be fixed in the covariance matrix of the difference vector $\tilde{u}_{km}$. However, in a conjoint choice experiment with several choice sets that may partially contain the same alternatives, the covariances of the differences $\tilde{u}_{km}$ are related between choice sets, which makes it difficult, if not impossible, to fix variance elements in the covariance matrix of these differences. Therefore, covariance parameters should be fixed in the original matrix Ω in conjoint choice experiments. When the X matrix is identical for all individuals so that there are no predictors that vary across subjects and all respondents receive the same choice sets, which often is the case in conjoint choice experiments, additional problems of model identification result (Heckman and Sedlacek 1985). However, a necessary[2] condition for identification of such a conjoint MNP model is that the total number of covariance parameters in Ω needs to be smaller than the total number of alternatives H.

A second problem of the general MNP formulation is that predictions for new profiles, not included in the conjoint design, cannot be made with the covariance matrix in (5) 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). To arrive at a model that is both identified and that allows for predictions of new profiles, we have to impose restrictions on Ω. The specification that enables the prediction of new alternatives that we propose assumes $\Sigma_e = I_H$, and we parameterize Σ_ψ as a matrix of rank one for reasons of parsimony and identification: $\Sigma_\psi = \sigma\,\sigma'$, with σ an S vector of parameters, where S is the number of columns in the X matrix. The number of parameters in Ω is then equal to the number of β parameters. A more general specification for Σ_ψ results in an increase in the number of covariance parameters so that identification often becomes a problem. Especially when the number of columns (S) in X or the number of profiles (H) is large, our specification for Ω is very parsimonious compared to a full random coefficients model or general Probit model.

In addition, we consider three possible structures for σ. The simplest is $\sigma = 0$, so

$$\Omega^{IP} = I_H, \tag{10}$$

and the model reduces to the well-known Independent Probit (IP) model, which is very similar to the MNL model and has similar properties, including IIA (Hausman and Wise 1978; Amemiya 1981). The IP model is nested in the second model we consider: a random intercepts model. Here only the intercepts (the coefficients of the dummies indicating the brand names) are taken to be random over individuals, but

[2]Because this is a necessary and not a sufficient condition, we check for local identification of the models in the applications by calculating the eigenvalues of the Hessian matrix in the optimum. When all eigenvalues are positive, this is a strong indication that the model is identified (based on Bekker et al. 1994). Note that more parameters can be estimated when the structure of the X-matrix is more general, empirical results indicate that if G groups of respondents each receive different choice sets, this leads to $G - 1$ additional degrees of freedom in estimation.

the other coefficients remain fixed. Let the brand dummies be collected in the submatrix X_1 of X, so $X = (X_1, X_2)$ and let σ be partitioned accordingly. Then

$$\Omega^{RI} = I_H + X_1\sigma_1\sigma_1'X_1'. \quad (11)$$

This model accounts for context effects for the brand names only because non-zero off-diagonal values in Ω are only related, through the σs, to the brand name attributes. A random intercepts MNL model has been applied previously by, for example, Gönül and Srinivasan (1993). The specification (11), denoted by RI, is nested in the third, most general model that we consider. Here we allow for random variation in all coefficients:

$$\Omega^{RC} = I_H + X\sigma\sigma'X'. \quad (12)$$

This random coefficients model is denoted by RC. It accounts for context effects caused by all attributes in the conjoint design. Rossi, McCulloch, and Allenby (1996) previously developed a random coefficients Bayesian MNL model. Our model differs from the models by Gönül and Srinivasan (1993) and Rossi, McCulloch, and Allenby (1996) in the factor structure on Σ_ψ and in being applied to conjoint choice experiments that have the typical structure of repeated choice sets of varying composition.

We now have a number of models to be considered. We can estimate the IP, RI, and RC models in two ways, each without and with assuming independence of the random utilities between choice sets, that is, using the likelihood in (6) or (9), respectively. We use IP, RI, and RC to denote the models that do not assume the choice sets independent, and attach a prefix "*i*" to denote the corresponding models when independence is assumed. This would result in six models, but in fact there are only five different models since the IP and iIP models are identical.[3] Table 1 gives an overview of those models.

[3]This follows because for the IP model we may freely set the variance of one element of e_j equal to zero, for example, the first element (the base alternative in our situation). The utility attached to that element then becomes nonrandom, and the composite probabilities pertaining to choice arrays simply factor out. For example, u_{j1} in (8) can freely be taken to be nonrandom. Hence, under IP the joint probability is the product of the probabilities for each choice set: $\pi_{24} = P(u_{j2} > u_{j1}, u_{j2} > u_{j3}) \cdot P(u_{j4} > u_{j1}, u_{j4} > u_{j5})$. Note that this property only holds when the choice sets have one alternative (the base alternative) in common.

Table 1 Overview of Models

	RC	RI	iRC	iRI	IP
Likelihood	L_2 (9)	L_2 (9)	L_1 (6)	L_1 (6)	L_1 (6)
Ω	$I_H + X\sigma\sigma'X'$	$I_H + X_1\sigma_1\sigma_1'X_1'$	$I_M + X\sigma\sigma'X'$	$I_M + X_1\sigma_1\sigma_1'X_1'$	I_M

X: dummy indicators for all attributes, X_1: dummy indicators for brands only.

To investigate the restrictiveness of the one-factor structure we compare our models with two special cases of the general random coefficients probit model (5). First, we consider an unrestricted probit covariance structure $\Omega = \Sigma_\epsilon$, and second we consider the full random coefficients structure $\Omega = I_M + X\Sigma_\psi X'$, both in the "independent" context of likelihood (6). Only the latter structure can be used to evaluate predictions. For both these models we impose the general necessary identification restrictions of MNP models.

The covariance parameters σ in (11) and (12) represent the variance in utility associated with each specific attribute level. According to the RC model, for example, a high estimate for a σ parameter contributes to a high covariance of alternatives sharing that feature, because element $\omega_{m,n}$ of Ω is equal to $X_m\sigma\sigma'X_n'$ for $m \neq n$. This is consistent with specific context effects (Simonson and Tversky 1992). The diagonal elements $\omega_{m,m}$ of Ω are equal to $1 + X_m\sigma\sigma'X_m'$. To interpret the estimates of σ one may inspect the estimated matrix $\hat{\Sigma}_\psi = \hat{\sigma}\hat{\sigma}'$ and/or the matrix $\hat{\Omega} = I + X\hat{\sigma}\hat{\sigma}'X'$. The significant values in the σ vector reveal possible sources of dependencies between the alternatives in the conjoint experiment and thereby enable inference on possible heterogeneity of attribute level coefficients due to context effects. The iRC and iRI models account for local tradeoff contrast effects, that is, similarity effects within choice sets, whereas the RC and RI models in addition account for background tradeoff contrast effects caused by attribute tradeoffs in other choice sets.

3. Applications

3.1. Introduction

The performance of the various models presented in the previous section is evaluated using the results from two conjoint choice experiments on cars and coffee makers. Sections 3.2 and 3.3 present the estimation results for the cars and coffee-maker data, respectively. In § 3.4 we compare, for the coffee-maker data, the performance of the models on holdout data and illustrate the simulation of market shares of new alternatives. Both studies used mall-intercept samples, attributes and levels were determined on the basis of in-depth interviews with manufacturers and consumers, and effects-type coding was used for the attribute dummies. For the simulations of the probabilities in the likelihood we use the SRC procedure of Hajivassiliou, McFadden, and Ruud (1993) with 100 draws. The optimization routine we use to maximize the likelihoods is the Broyden, Fletcher, Goldfarb, and Shanno algorithm implemented in the Gauss-package (Aptech 1995). All models were estimated on a 133-MHz Pentium PC.

All models were started with all parameters equal to zero (this means that the probabilities in each choice set are equal to $1/M$ at the start of the estimation). The likelihood of the IP model has a unique maximum (Maddala 1983), and convergence is quick in general. Other MNP covariance structures can cause convergence to local optima (e.g., Daganzo 1979; McCulloch and Rossi 1994). Therefore, we also started the RC model from the IP estimates for β (with $\sigma = 0$) as well as from 10 sets of random starting values for all parameters. These yielded virtually the same results and will not be reported. This supports the findings of McCulloch and Rossi (1994), who also found no evidence of a multimodal likelihood for the MNP model.

To compare the models we use the log-likelihood value, AIC (Akaike 1973), BIC (Schwarz 1978), and the Pseudo R^2 value (e.g., McFadden 1976) with a null model in which all probabilities in a choice set are equal to $1/M$. Furthermore,we use the likelihood ratio test to compare the nested models. Because the RC model and the iRC model are non-nested and the RI and iRI models are non-nested, we cannot test between those models with the LR test. But, because they do have the same number of parameters, they can easily be compared on the other statistics.

3.2. Estimation Result for the Car Data

In the car experiment, respondents had to choose from nine choice sets, each with four alternatives. The attributes and levels of the cars are listed in Table 2. The last alternative in each choice set was the "no-choice" option, which was used as the base alternative. There were 398 respondents, divided into six groups that received different choice sets.[4] The variables Price, Fuel Consumption, and Engine Capacity were specified to be linear with codes -1, 0, 1 for the levels, respectively.

Table 3 gives the estimation results for all five models.[5] All eigenvalues of the final Hessian are positive for all models, which is a strong indication that all models are identified (Bekker et al. 1994).[6] The log-likelihoods show that, when we relax only the independence between alternatives (the IIA property), a significantly better fit results. This holds for the iRI model ($LR(9) = 27.7$, $p < 0.01$) and the iRC model ($LR(15) = 40.0$, $p < 0.01$) relative to the IP model. However, the iRC model does not fit significantly better than the iRI model ($LR(6) = 12.2$, $p > 0.05$). Note that although the iRI model has a significantly better fit than the IP model, none of the covariance parameters is significant. When we also relax the independence between choice sets we see a further substantial improvement in the log-likelihoods for the RC and RI models (which occurs without the cost of increasing the number of parameters). In both cases the improvement is significant ($LR(9) = 601.2$, $p < 0.01$, for the RI model and $LR(15) = 626.4$, $p < 0.01$, for the RC model). Furthermore, the RC model is significantly better than

[4]In this design, with 27 profiles and six groups, more than 27 parameters can be estimated. We estimated this data set for one group, two groups up to all six groups of respondents. The number of positive eigenvalues for each of those analyses provide empirical evidence that each additional group of respondents results in one extra parameter that can be estimated up to the maximum as specified by the standard identification rule in MNP models (see e.g., Keane 1992). This would imply that for the complete data set with six groups, only two degrees of freedom are left. Note, however, that although when parameters are formally identified, nonpositive eigenvalues may be obtained due to data limitations.

[5]For this data set we do not compare the results of the models in Table 1 with the general Probit and Full Random Coefficients structure for the covariance matrix, because the number of parameters in these models are excessively large: 143 covariance parameters for the general Probit model and 119 for the Full Random Coefficients structure. Furthermore, we do not have holdout sets for this data set, so we cannot compare models on predictive performance.

[6]The estimation times range from less than 20 minutes for the IP model up to 44 hours for the RC model.

Table 2 Attributes and Levels of Cars

Attribute Level	Brand	Price (Dfl)	Fuel Consumption (L/100km)	Engine Capacity	Power Steering	Doors	Airbag
1	Renault 19	27.000, –	5.3	1.4 l.	Yes	2/3	Yes
2	Alfa Romeo 33	30.000, –	6.3	1.6 l.	No	4/5	No
3	Opel Vectra	33.000, –	7.7	1.8 l.			
4	VW Golf						
5	Volvo 440						
6	Daihatsu Applause						
7	Ford Escort						
8	Nissan Sunny						
9	Kia-Sephia SLX						

the RI model (LR(6) = 25.2, $p < 0.01$). The AIC and BIC statistics support the RC and RI models as the best models. The BIC statistic, which imposes a more severe penalty, favors the IP model over the iRC and iRI models and the RI model over the RC model. Note that the number of observations that appears in the BIC statistic is equal to the number of respondents for the RC and RI model, while it is equal to K (the number of choice sets) times the number of respondents for the iRC, iRI, and IP models.

From the value of the log-likelihood it is obvious that the RC model fits much better than the iRC model. The increased fit comes at no cost of additional parameters to be estimated. The same holds for the RI and iRI models. This is our most important finding: The model fit improves substantially when choice sets are not treated as independent. Table 3 also shows that the RI model provides a much better fit than the iRC model, with a much smaller number of parameters. Apparently, it is crucial to account for heterogeneity in the attribute levels across choice sets, which may indicate the presence of strong background tradeoff contrast effect in this choice experiment.

The signs of the regression parameters are as expected for all attributes. The differences between the models are not large in most cases. Subjects prefer lower price, higher mileage, more engine capacity, power steering, fewer doors (smaller cars), and an airbag. The brand intercepts indicate some differences in preferences among brands. Furthermore, Table 3 shows that the standard errors of the parameters are almost always higher in the iRC and iRI model compared to the RC and RI model, which indicates that these former models are misspecified.

The estimates for the covariance parameters σ lead to variances in the RC model ranging from 1.00 (for the base alternative) to 3.90, while the covariances are all positive and range from 0.00 (for the base alternative) to 2.59. Similar observations can be made for the covariance matrix of the RI and iRC models, where the iRC model is the only model that has negative covariances. It is of some interest to compare the estimates of σ for the RC and iRC models. In the RC model, the σ estimates for the brand dummies are much larger than in the iRC model, which may indicate that the background contrast effect predominantly involves brands. That is, the extent to which a brand has a high utility depends on the tradeoff between brands made in other choice sets. Note that this is supported by comparison of the RI and iRI models.

3.3. Estimation Results for the Coffee-Maker Data

The five attributes for the coffee-makers—brand name, capacity, price, presence of a special filter, and thermos-flask—are listed in Table 4. Using a factorial design, 16 profiles were constructed. Data were collected from 185 respondents, divided into two groups that received different choice sets based on the same 16 profiles. Respondents had to choose from eight sets of three alternatives and from four sets of five alternatives. Each choice set included the same base alternative (given this design, there is only one degree of

Table 3 Estimation Results Car Data

Parameter	RC	RI	iRC	iRI	IP
β_1 Renault	0.690* (.104)	0.687* (0.89)	0.166 (.116)	0.117 (.187)	0.205* (.056)
β_2 Alfa	0.767* (.114)	0.761* (.099)	0.227 (.115)	0.245 (.132)	0.288* (0.57)
β_3 Opel	1.217* (.114)	1.239* (.100)	0.847* (.103)	0.764* (.093)	0.723* (.057)
β_4 VW	1.253* (.108)	1.244* (.094)	0.826* (.080)	0.795* (.070)	0.751* (.052)
β_5 Volvo	0.794* (.112)	0.792* (.100)	0.335* (.082)	0.209 (.138)	0.316* (.057)
β_6 Daihatsu	−0.079 (.120)	−0.072 (.102)	−0.818* (.188)	−1.112* (.506)	−0.549* (.066)
β_7 Ford	0.511* (.109)	0.507* (.096)	−0.124 (.147)	−0.026 (.127)	0.029 (.061)
β_8 Nissan	0.112* (.126)	0.139 (.106)	−0.344* (.098)	−0.307 (.208)	−0.248* (.061)
β_9 Kia-Sephia	−0.426* (.125)	−0.420* (.110)	−1.119* (.163)	−0.882* (.123)	−0.851* (.072)
β_{10} Price	−0.066* (.023)	−0.061* (.022)	−0.059 (.038)	−0.058* (.028)	−0.059* (.021)
β_{11} Fuel consumption	0.103 (.023)	0.094 (.022)	0.093* (.034)	0.095* (.033)	0.092* (.021)
β_{12} Engine capacity	0.197* (.023)	0.195 (.022)	0.265* (.041)	0.208* (.030)	0.180* (.021)
β_{13} Power steering	0.046* (.020)	0.049* (.020)	0.057* (.027)	0.051* (.022)	0.052* (.019)
β_{14} Number of doors	−0.211* (.020)	−0.199* (.019)	−0.184* (.036)	−0.198* (.024)	−0.173* (.018)
β_{15} Airbag	0.045* (.019)	0.050* (.019)	0.066 (.038)	0.068* (.022)	0.055* (.018)
σ_1 Renault	1.145* (.117)	1.157* (.103)	−0.386 (.434)	−0.893 (.509)	
σ_2 Alfa	1.384* (.122)	1.404* (.113)	0.284 (.423)	0.379 (.515)	
σ_3 Opel	1.345* (.117)	1.422* (.107)	1.132* (.562)	−0.682 (.556)	
σ_4 VW	1.288* (.110)	1.339 (.102)	0.331 (.371)	−0.296 (.512)	
σ_5 Volvo	1.357* (.121)	1.432* (.112)	0.069 (.422)	1.261 (.659)	
σ_6 Daihatsu	1.173* (.144)	1.126* (.128)	−0.779* (.319)	1.028 (.600)	
σ_7 Ford	1.294* (.120)	1.330* (.111)	−0.681* (.292)	0.299 (.225)	
σ_8 Nissan	1.567* (.151)	1.508* (.132)	0.119 (.620)	0.396 (.741)	
σ_9 Kia-Sephia	1.367* (.145)	1.289* (.135)	0.505 (.472)	−0.054 (1.06)	
σ_{10} Price	−0.015 (.036)		0.221* (.106)		
σ_{11} Fuel consumption	−0.009 (.029)		−0.097 (.108)		
σ_{12} Engine capacity	−0.023 (.034)		0.070 (.090)		
σ_{13} Power steering	−0.001 (.020)		0.183* (.069)		
σ_{14} Number of doors	0.135* (.027)		−0.023 (.075)		
σ_{15} Airbag	0.043 (.029)		0.346* (.134)		
Log-likelihood	−4114.257	−4126.859	−4407.485	−4413.598	−4427.468
AIC	8288.513	8301.719	8874.970	8875.196	8884.936
BIC	8408.108	8397.393	9060.480	9023.604	8977.691
Pseudo R^2	0.172	0.169	0.112	0.111	0.108

*$p < 0.05$, standard errors in parentheses.

Table 4 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, –		

freedom left in the iRC model). Furthermore, eight holdout profiles were constructed; four holdout sets with three alternatives and two holdout sets with five alternatives, where the same base alternative was used as in the estimation data. These holdout sets were offered to all respondents. The choices from the sets with three alternatives and the choices from the sets with five alternatives are modeled separately.

In Table 5 the parameter estimates of all models, as

Table 5 Estimation Results for Coffee-Maker Data

	Choice sets (8) with three alternatives					Choice sets (4) with five alternatives				
Attribute (level)	RC	RI	iRC	iRI	IP	RC	RI	iRC	iRI	IP
β_1 Brand (1)	−0.029 (.101)	−0.050 (.081)	−0.106 (.203)	−0.100 (.077)	0.015 (.055)	0.063 (.060)	−0.053 (.066)	−0.081 (.118)	0.116 (.097)	0.019 (.046)
β_2 Brand (2)	−0.240* (.078)	−0.224* (.068)	−0.179 (.222)	−0.137 (.072)	−0.265* (.047)	−0.314* (.065)	−0.189* (.051)	−0.140 (.084)	−0.037 (.098)	−0.210* (.043)
β_3 Capacity (1)	−1.075* (.092)	−0.912* (.052)	−1.166* (.123)	−1.088* (.100)	−0.778* (.050)	−0.943* (.105)	−0.639* (.054)	−1.303* (.261)	−0.831* (.087)	−0.597* (.051)
β_4 Capacity (2)	0.565* (.060)	0.431* (.042)	0.587* (.090)	0.504* (.058)	0.372* (.039)	0.441* (.052)	0.258* (.040)	0.570* (.103)	0.322* (.051)	0.236* (.038)
β_5 Price (1)	0.432* (.116)	0.212* (.071)	0.326 (.497)	0.275* (.081)	0.217* (.065)	0.282* (.075)	0.238* (.050)	0.983 (.585)	0.285* (.062)	0.213* (.048)
β_6 Price (2)	0.244* (.082)	0.333* (.052)	0.378 (.269)	0.407* (.064)	0.296* (.048)	−0.001 (.096)	0.081 (.041)	0.760 (.636)	0.178* (.053)	0.098* (.040)
β_7 Filter (1)	0.355* (.038)	0.294* (.031)	0.354* (.098)	0.372* (.047)	0.261* (.029)	0.232* (.039)	0.211* (.032)	0.257* (.056)	0.342* (.051)	0.209* (.031)
β_8 Thermos (1)	0.393* (.054)	0.253* (.035)	0.269 (.152)	0.197* (.041)	0.244* (.034)	0.359* (.048)	0.280* (.035)	0.400* (.080)	0.258* (.042)	0.273* (.033)
σ_1 Brand (1)	0.417* (.096)	0.787* (.068)	0.717 (.386)	1.096* (.206)		0.133 (.078)	0.564* (.075)	0.529* (.209)	0.849* (.260)	
σ_2 Brand (2)	−0.387* (.099)	−0.574* (.077)	0.107 (.665)	−0.518* (.197)		−0.075 (.075)	−0.285* (.081)	0.185 (.201)	0.484* (.186)	
σ_3 Capacity (1)	0.850* (.094)		0.585* (.270)			0.803* (.106)		0.606* (.175)		
σ_4 Capacity (2)	−0.348* (.083)		−0.001 (.224)			−0.052 (.075)		−0.239 (.229)		
σ_5 Price (1)	−0.562* (.139)		−0.084 (.599)			−0.541* (.108)		1.520 (.899)		
σ_6 Price (2)	−0.145 (.100)		0.482 (.473)			0.034 (.076)		1.173 (.709)		
σ_7 Filter (1)	0.023 (.058)		0.298 (.372)			0.070 (.064)		0.338* (.133)		
σ_8 Thermos (1)	−0.206* (.071)		0.173 (.259)			−0.238* (.064)		−0.369* (.110)		
Log-likelihood	−1086.622	−1192.833	−1279.100	−1288.123	−1299.897	−934.589	−993.781	−997.026	−1005.647	−1014.191
AIC	2205.245	2405.666	2590.201	2596.245	2615.793	1901.178	2007.561	2026.053	2031.294	2044.382
BIC	2256.770	2437.870	2674.997	2649.244	2658.192	1952.704	2039.766	2099.758	2077.361	2081.235
Pseudo R^2	0.332	0.266	0.213	0.208	0.201	0.215	0.166	0.163	0.156	0.148

$^*p < 0.05$, standard errors in parentheses.

well as the statistics for model comparison, are listed for the three-alternatives estimation data and for the five-alternatives estimation data. All eigenvalues of the Hessian are positive for all models, indicating that they are identified (Bekker et al. 1994).[7]

Table 5 shows that the RC model has by far the highest value for the log-likelihood, both for the three- and five-alternatives data. The LR statistic for testing the RC model against the IP model is significant (LR(8) = 426.6, $p < 0.01$, for the three-alternatives data and LR(8) = 159.2, $p < 0.01$, for the five-alternatives data). The iRC model also has a significantly better value for the log-likelihood than the IP model in both situations (LR(8) = 41.6, $p < 0.01$, and LR(8) = 34.3, $p < 0.01$, respectively). Similar results are found for the RI and iRI models (LR(2) tests, $p < 0.01$). Both for the three- and five-alternatives results the RC specification fits significantly better than the RI specification and the iRC specification fits significantly better than the iRI specification (LR(6) tests, $p < 0.01$). From the value of the log-likelihood it is obvious that the RC model, which accounts for correlations between choice sets, fits much better than the iRC model, which treats all choice sets as independent. The number of parameters for these two models is equal, so that the increased fit comes at no cost of additional parameters estimated. Interestingly, the RI model provides a much better fit than the iRC model with a much smaller number of parameters. This holds both for the three- and five-alternatives data. Again it is important to account for correlations across choice sets. Furthermore, allowing for random variation in all β parameters leads to a significantly better fit of the RC models compared with the RI models that only have random coefficients for the brand intercepts (cf. Rossi and Allenby 1993). The AIC and BIC criteria support the RC model as the best

[7]The estimation times range from less than a minute for the IP models up to 4–6 hours for the RC models.

model and the RI model as the second best. Again in this application, our most important finding is that the model fit improves substantially when the choice sets are not treated as independent, which holds for both the RC and RI specifications. This may again indicate the presence of rather strong background tradeoff contrast effects.

For the sets with three and five alternatives we also compare the results of the five models in Table 1 with the full covariance Probit model and a Full Random Coefficients structure to investigate the restrictiveness of the one-factor structure imposed on the covariance matrix. For the choice sets with three alternatives, the full covariance probit model has a likelihood of -1275.125 ($R^2 = 0.216$) and the Full Random Coefficients model has a likelihood of -1271.189 ($R^2 = 0.218$). For the sets with five alternatives, the full covariance probit model has a likelihood of -986.729 ($R^2 = 0.172$) and the Full Random Coefficients model has a likelihood of -989.035 ($R^2 = 0.170$). Table 5 shows that the likelihoods of these two models are higher than those of the other models, except for the RC model (and by the RI model for the sets with three alternatives), which shows that the one-factor RC structure provides a better fit with a much more parsimonious representation.

In Table 5, the signs of the regression parameters are as expected for all models; the lowest capacity and the highest price have a negative partial utility and the attributes Thermos-Flask and Special Filter have a positive partial utility when present. The estimates of the regression parameters for the five-alternatives data are similar to the three-alternatives data for each model in general. From Table 5 it can be observed that due to misspecification, by assuming that choice sets are independent, there is a loss of statistical efficiency leading to larger standard errors both for the iRC and iRI specifications compared to the RC and RI specifications, respectively.

The RC model has six significant covariance parameters for the three-alternatives data and three significant covariance parameters for the five-alternatives data. These covariance parameters are responsible for the large increase in model fit of the RC model over the IP model since differences in β estimates are modest. This is in line with similar findings by Börsch-Supan et al. (1990). Although, on the basis of the β estimates, all models predict the highest expected utility for the same profile, due to the differences in the estimated covariance structure, the predicted choice probabilities differ between models.

The variances of the profiles, calculated from the estimates of the RC model σ parameters, range from 1.00 to 3.26 for the three-alternatives data and from 1.00 to 2.22 for the five-alternatives data. The covariances range from -2.21 to 1.73 for the three-alternatives data and from -1.22 to 1.15 for the five-alternatives data. Interestingly, the covariance matrix of the RC model for the three- and five-alternatives choice sets data reveal alternatives with near zero covariances with all other alternatives. This indicates that these are (almost) independent of the other alternatives. Although the superior fit of the iRC and iRI models over the IP model may be indicative of local contrast effects, due to which utilities within choice sets are correlated and the IIA property does not hold, the superior fit of the RC and RI models may indicate again that background contrast effect are prevalent.

A comparison of the estimates for σ in the RC model for the three-alternatives data with those for the five-alternatives data reveals that most of the latter are smaller in absolute sense. This may indicate that the background contrast effect is larger for the three-alternatives data. This can be explained from the observation that the range of attribute levels within choice sets is necessary smaller when there are three alternatives in the choice sets, while the number of choice sets is larger, thus giving much more latitude for background contrast effect to occur. An interesting finding is that for the iRC model the reverse seems to hold. The estimates for σ tend to be larger for the five-alternatives data than for the three-alternatives data. The iRC model for the three-alternatives data even has a covariance matrix close to the identity matrix because most of the σ estimates are insignificant. This may occur because local contrast effects within each choice set are much more prevalent for larger choice sets than for smaller choice sets. Thus, the decision of the size of the choice set in conjoint choice experiments seems to affect the balance of local versus background contrast effects, where an increasing choice set causes a shift from background to local contrast effects.

3.4. Holdout Predictions and Market-Share Simulations for the Coffee-Maker Data

The estimates of the models on the three- and five-alternatives coffee-maker data were used to predict the holdout sets with three and five alternatives, respectively. Table 6 gives the log-likelihood, pseudo R^2, and AIC and BIC values for the predicted choices in the holdout choice sets. It shows all statistics indicate that the RC model predicts the holdout sets much better compared with the other models in both the three- and five-alternatives holdout sets. The LR statistic for testing this model against the IP model is significant in both cases (LR(8) = 160.6, $p < 0.01$, and LR(8) = 58.5, $p < 0.01$, respectively). The iRC and iRI models do not predict the holdout sets better than the IP model, despite the fact that the fit (Table 5) was significantly better. The RI model only predicts significantly better than the IP model for the holdout sets with three alternatives (LR(2) test, $p < 0.01$) and it predicts better than the iRC model with a smaller number of parameters. In both situations the RC model predicts significantly better than the RI model (LR(6) test, $p < 0.01$).

To further investigate the restrictiveness of the one-factor covariance structure we compare the results with those of the Full Random Coefficients model (note that the full covariance Probit model cannot be used to generate predictions in a conjoint choice context). The Full Random Coefficients model has for the holdout sets with three alternatives a predicted likelihood of -750.884 ($R^2 = 0.076$) and for the holdout sets with five alternatives a predicted likelihood of -506.411 ($R^2 = 0.150$). Table 6 shows for the holdout sets with three alternatives that this model is again outperformed with respect to the likelihood by the RC model. For the holdout sets with five alternatives it is outperformed by all models.

Table 6 shows that the prediction of holdout sets with five alternatives is much better than the prediction of holdout sets with three alternatives. The pseudo R^2 value is 0.165 for the RC model predictions for holdout sets with three alternatives, and it is equal to 0.296 for the RC model predictions for holdout sets with five alternatives. The same pattern holds for the other models. This may be caused by the relative importance of local and background contrast effects. The five-alternatives RC model estimates indicated more local but less background contrast effects. The background contrast effects will have a different influence in the holdout task, since the range of choice sets offered in the holdout task differ substantially from that in the main task. However, this effect will be stronger for the holdout task with three alternatives than for the holdout task with five alternatives. This, therefore, reduces the predictive validity of the estimates from the calibration sample for the three-alternatives task relatively more. In addition, the within-choice set variation in attributes is larger for the five- than for the three-alternatives tasks, due to which the differences in within-choice set variation between the calibration and holdout task are less for the five-choice set task. This enhances similarity of local tradeoff contrast effects in the holdout tasks, resulting in a better predictive validity of the five-choice set task.

This leads us to conclude that larger choice sets may be preferred in conjoint choice experiments to alleviate background contrast effects and obtain higher predictive validity. The above analyses show that not only

Table 6 Holdout Predictions for Coffee-Maker Data

	Holdout sets (4) with three alternatives					Holdout sets (2) with five alternatives				
	RC	RI	iRC	iRI	IP	RC	RI	iRC	iRI	IP
Log-likelihood	−679.075	−750.198	−784.677	−779.071	−759.368	−419.302	−447.012	−488.336	−464.750	−448.444
AIC	1390.151	1520.395	1601.354	1578.142	1534.736	870.604	914.023	1008.672	949.500	912.888
BIC	1441.676	1552.600	1675.060	1624.209	1571.589	922.130	946.228	1071.288	988.635	944.196
Pseudo R^2	0.165	0.077	0.035	0.042	0.066	0.296	0.249	0.180	0.220	0.247

the RC model performs much better on model fit but that also the holdout predictions are significantly better. This underlines the importance of considering correlations across choice sets to account for background contrast effects.

Now we illustrate that the RC model can lead to substantially different predictions of market shares than the IP and the iRC models. We consider three managerially relevant situations: a product modification, a product line extension, and the introduction of a me-too brand. In our simple hypothetical illutration we use the four profiles listed in Table 7.

We use the estimates obtained from the three-alternatives data from Table 5. First consider the situation of a product modification of the brand Philips. We assume that the current market consists of two products: Philips (profile 1) and Braun (profile 2). Assume that Philips modifies its existing product by introducing a thermos-flask and asking a higher price for the product (profile 4). The IP model predicts an increase in market share of 16.8%, while the RC model predicts an increase of only 11.5% in market share of Philips (Table 8). In the IP model the market shares of the two brands reverse approximately compared to the initial situation. In the RC model the market shares change less compared to the IP model. The iRC model is in-between the two other models with respect to market shares. However, compared with the "before" situation, the differences in market shares are the highest for this model.

The second example pertains to a product line extension. Assume that Philips modifies the existing brand and introduces it as an extension of its product line. The product (profile 4) differs from the existing product (profile 1) in that it has as an additional feature (a thermos-flask) and that it has a higher price. The market shares predicted by the IP, iRC, and RC models, before and after the product line extension, are provided in Table 8. The RC and IP models predict almost the same market share in the initial situation with two products. The iRC model predicts slightly different initial market shares. After the product line extension, the IP model predicts that the market share of Braun drops by 26.2%, whereas the RC model predicts only a 19.3% decrease of its market share. The RC model predicts a lower market share for the new Philips product compared with the IP model, and it predicts that this market share is drawn relatively more from the existing product of Philips. The market share of Philips is 67.3% as predicted by the IP model and 61.6% for the RC model. When two alternatives are similar (as are the two products of the same brand), the IP model predicts a too high joint probability (market share) of these two alternatives, due to the IIA property (Green and Srinivasan 1978). Note that the market share of the new Philips product predicted by the iRC model is higher than for both the RC and IP models.

The third situation considered is the introduction of a new brand with characteristics relatively similar to those of (one of) the brands already in the market, a me-too brand. Consider again the situation of two brands in the market (Philips and Braun). Now a third brand, Moulinex, introduces a coffee maker close to the existing product of the initial market-leader Braun. Table 8 gives the market shares predicted by the RC, iRC, and IP models before and after the introduction. Table 8 shows that the predictions of the models are quite different. After the introduction, the IP and iRC models predict that the new brand becomes the market leader, whereas the RC model still predicts the highest market share for the initial market-leader Braun and the lowest market share for the new brand. The RC model predicts that the initial market leader loses 13.5% market share as a result of the introduction of the me-too brand, while the IP model predicts it to lose 21.8%, and the iRC model even predicts a loss of 26.5%. In each case Philips loses around 14%. Obviously, the misspecification of the iRC and IP models results in an erroneous prediction that the new brand becomes the market leader. The RC model predicts that it will stay well behind the initial market leader and that it will be the smallest brand. The simpler iRC and IP models yield

Table 7 Attributes of Prediction Profiles

Attribute Profile	Brand	Capacity	Price (Dfl)	Special Filter	Thermos-Flask
1	Philips	10 cups	39, –	No	No
2	Braun	15 cups	69, –	No	Yes
3	Moulinex	15 cups	69, –	No	No
4	Philips	10 cups	69, –	Yes	No

Table 8 Market Simulations

	Product Modification				Product Line Extension				Me-too Brand			
	Brand	RC	iRC	IP	Brand	RC	iRC	IP	Brand	RC	iRC	IP
Before	P (1)	0.423	0.362	0.410	P (1)	0.423	0.362	0.410	P (1)	0.423	0.362	0.410
	B (2)	0.577	0.638	0.590	B (2)	0.577	0.638	0.590	B (2)	0.577	0.638	0.590
After	P (4)	0.538	0.549	0.578	P (1)	0.182	0.155	0.213	P (1)	0.283	0.226	0.244
	B (2)	0.462	0.451	0.422	B (2)	0.384	0.361	0.328	B (2)	0.442	0.373	0.372
					P (4)	0.434	0.484	0.460	M (3)	0.276	0.401	0.384

incorrect predictions because they do not accommodate effects of the covariance structure of utilities within and between choice sets. It is apparent that this misspecification may severely affect market simulations, even if the estimates of the attribute parameters of these models are relatively similar.

4. Conclusion and Discussion

In this paper we applied the MNP model for the analysis of conjoint choice experiments. The MNP model does not treat choices coming from the same respondent as independent, but can be used to model correlations of choice alternatives within and between choice sets. We have taken two hurdles in the application of the MNP model for conjoint choice data, i.e., that of model identification and that of making predictions. The resulting RC model is much more parsimonious than the full MNP or random taste variation MNP models and outperforms those in terms of fit and predictive validity. Given the identification problems of MNP models, we recommend that in applications of MNP models to conjoint choice data identification of the model is thoroughly investigated, for example, by checking the necessary condition that the number of parameters should be smaller than the number of profiles, as well as the eigenvalues of the Hessian in the optimum.

The proposed random coefficients MNP model accounts for heterogeneity due to context effects in modeling stated choices in experimental choice analysis. In addition, it does not suffer from the restrictive IIA property. We showed that the proposed model leads to significantly better fit and predictions than random intercepts and IP models. The iRC model that arises as a special case of the RC model by assuming that choice sets are independent does not suffer from the IIA property, but does not account for correlations between choice sets. This model has a significantly better fit than the IP model, but does not predict holdout sets better. An interesting finding is that a random intercepts model, where only brand intercepts are assumed heterogeneous across subjects and choice sets, provides a better fit and holdout predictions than the iRC model with fewer parameters. The RC model is significantly better in terms of model fit and predicts the holdout sets significantly better compared with all other models. Therefore, we conclude that it is important to allow all coefficients to be random, but that it is even more important to take account of correlations between choice sets. An alternative approach by which this can be accomplished is the Bayesian approach by, for example, Rossi and Allenby (1993). This procedure provides the advantages of allowing for inferences on part-worth estimates at the respondent level. A comparison of the SML and Bayesian approaches in terms of their performance is left for further research.

We hypothesized that the covariances among utilities within a choice set (as modeled in the iRC and iRI models) account for local tradeoff contrast effects (Simonson and Tversky 1992). Such local contrast effects specify that a respondents' tradeoff of two alternatives in a specific choice set is affected by other tradeoffs in that choice set. These effects would be reflected in significant covariance parameters in the iRC and iRI models, which were observed in both applications. Thus, our results are indicative of such local contrast effects, but it must be noted that the estimates of our model cannot be considered as unambiguous evidence because other factors, including heterogeneity, may also have caused such covariances.

The superior fit of the RC and RI models in both applications is caused by accommodating covariances in utilities among choice sets, which may be indicative of background tradeoff contrast effects (Simonson and Tversky 1992). Such background contrast effects occur when tradeoffs made among profiles in other choice sets affect the tradeoffs in the current choice sets. Our results strongly indicate the existence of such background contrast effects.

Judged from the magnitudes of the relevant covariance parameters estimates, the background contrast effect may vary in importance for different attributes. Such variation may be explained from different types of context effects, such as substitution, dominance, attraction, and compromise effects. Further research should investigate when these different types of effects occur in conjoint choice experiments and how they are reflected in the covariance parameters in our model. The results showed that the background contrast effect may be stronger for smaller choice sets due to the more restricted range of attribute levels in such sets, whereby previous tradeoffs therefore may have a larger effect. Since predictions for holdout sets deteriorated for smaller holdout choice sets, such background contrast effects may negatively affect the external validity of conjoint choice experiments. We recommend that experiments are conducted to investigate context effects in conjoint choice designs while future research should delve into the question of optimal design of conjoint choice experiments in terms of the number, size, and composition of choice sets in view of minimizing context effect.[8]

Appendix: The SRC Probability Simulator

The multidimensional integrals involved in the RC model cannot be evaluated numerically. However, simulation techniques can be used to approximate the integrals in (4). The simulators differ basically as to the way the drawings from the error distribution are obtained. Hajivassiliou, McFadden, and Ruud (1993) compared the known simulators and concluded that the Smooth Recursive Conditioning (SRC) simulator, also known as the GHK simulator (after Geweke, Hajivassiliou, and Keane), is one of the best.

This section is based on Börsch-Supan and Hajivassiliou (1993), with some different notation to make this appendix consistent with the earlier notation (for the SRC simulator, see also, e.g., Hajivassiliou 1993, and Geweke, Keane, and Runkle 1994).

We start with the following discrete choice model with one choice set (hence $K = 1$, and we therefore suppress this index):

$$Y^* = X\beta + \epsilon = \mu + \Gamma e,$$

with

$$Y^* \sim N_M(\mu, \Omega), \quad \Omega = \Gamma\Gamma',$$
$$e \sim N_M(0, I), \quad \epsilon = \Gamma e. \tag{A1}$$

Observed is the $(M \times 1)$ vector Y, where

$$y_m = \begin{cases} 1 & \text{if } y_m^* \geq y_n^* \ \forall \ n = 1, \ldots, M, \ n \neq m, \\ 0 & \text{other} \end{cases}. \tag{A2}$$

The probability $P(y_m = 1)$ involves an $(M - 1)$-dimensional integral that cannot be evaluated for $M > 4$. The SRC probability simulator simulates the probability $\pi_m = P(y_m = 1)$ as follows:

Let V be a nonsingular matrix and define L as the lower Choleski-factor of $V\Omega V' = LL'$. Then from (A1) and (A2) it follows that for all $n \neq m$ it must hold that:

$$\begin{aligned} -\infty &\leq y_n^* - y_m^* \leq 0, \\ \Rightarrow -\infty &\leq VY^* \leq 0, \\ \Rightarrow -\infty &\leq V(X\beta + \epsilon) \leq 0, \\ \Rightarrow -\infty &\leq V(\mu + \Gamma e) \leq 0, \\ \Rightarrow -\infty - V\mu &\leq Le \leq 0 - V\mu, \\ \Rightarrow a &\leq Le \leq b. \end{aligned} \tag{A3}$$

We now need random drawings such that:

$$\begin{aligned} &e_1 \sim N(0,1) \text{ s.t. } a_1 \leq l_{11}e_1 \leq b_1, \\ &e_2 \sim N(0,1) \text{ s.t. } a_2 \leq l_{21}e_1 + l_{22}e_2 \leq b_2, \\ &\quad\vdots \qquad\qquad\qquad \vdots \\ &e_M \sim N(0,1) \text{ s.t. } a_M \leq l_{M1}e_1 + \cdots + l_{MM}e_M \leq b_M, \\ &\qquad \Leftrightarrow a_M \leq L_{M,<M}e_{<M} + l_{MM}e_M \leq b_M. \end{aligned} \tag{A4}$$

These drawings can be obtained by drawing a vector $U \sim$ Uniform(0,1) and by calculating from this vector U:

$$e = \Phi^{-1}\left[(\Phi(b^*) - \Phi(a^*))U + \Phi(a^*)\right], \tag{A5}$$

where Φ is the normal cumulative distribution function. Then it holds that e follows a truncated normal distribution (Hajivassiliou and McFadden 1990):

$$e \sim N(0,1) \text{ s.t. } a^* \leq e \leq b^*, \quad \text{with } -\infty \leq a^* \leq b^* \leq \infty. \tag{A6}$$

Now define Q_m $(m = 1, \ldots, M)$ as:

$$\begin{aligned} Q_1 &= P\left(\frac{a_1}{l_{11}} \leq e_1 \leq \frac{b_1}{l_{11}}\right) \\ &\vdots \\ Q_m(e_1, \ldots, e_{m-1}) &= P\left(\frac{a_m - L_{m,<m}e_{-m}}{l_{mm}} \leq e_m \leq \frac{b_m - L_{m,<m}e_{<m}}{l_{mm}}\right). \end{aligned} \tag{A7}$$

[8]This paper has greatly benefitted from the suggestions of the editor Richard Staelin, the area editor Greg Allenby, and the anonymous reviewers. We also thank Peter Rossi for his insightful remarks.

The SRC probability simulator for π_m is then for random drawings $e_r = (e_{1r}, \ldots, e_{Mr})'$, and with R replications ($r = 1, \ldots, R$), defined as:

$$\pi_m^{SRC} = f_m(\beta,\Omega) = \frac{1}{R}\sum_{r=1}^{R}\prod_{m=1}^{M} Q_m\ (e_{1r}, \ldots, e_{m-1,r}) \qquad \text{(A8)}$$

The simulator is unbiased and smooth. That is, π_m^{SRC} is a continuous and differentiable function of β and Ω. The extension of the above to the conjoint situation with $K > 1$ and probabilities as in (5) is straightforward.

Estimation by Simulated Maximum Likelihood

With the method of Simulated Maximum Likelihood (SML), only the probabilities of the selected alternatives have to be simulated, which is computationally efficient. This method is known as Smooth SML (SSML) when it is applied with a smooth choice simulator as the SRC simulator. With SSML asymptotically efficiency requires that $R/\sqrt{J} \to \infty$ as $J \to \infty$ (Börsch-Supan and Hajivassiliou 1993), where R is the number of simulations. However, several studies show that SSML is efficient even when the number of simulations is rather low, say 10 to 20 (Mühleisen 1991; Lee 1992; Börsch-Supan and Hajivassiliou 1993; Geweke, Keane, and Runkle 1994). The simulated probabilities replace the probabilities π_{km} in the likelihood of (6) or π_l in the likelihood of (9).

A potential drawback of using smooth simulators is that the simulated probabilities are not restricted to add up to one over the M choices (McFadden 1989; Mühleisen 1991; Lee 1992). Lee (1992) stated that the adding-up property can always be satisfied by normalizing the original simulators, at the extra cost of simulating choice probabilities for all alternatives.[9]

[9]In the RC and RI model we do not use Lee's (1992) procedure because we then would have to simulate all M^K probabilities, which is in general not feasible within a reasonable amount of time. However, we expect that this will not be a problem due to the high accuracy of the SRC simulator.

References

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

Allenby, G. M., N. Arora, J. L. Ginter 1995. Incorporating prior knowledge into the analysis of conjoint studies. *J. Marketing Res.* 32 152–162.

——, J. L. Ginter 1995. Using extremes to design products and segment markets. *J. Marketing Res.* 32 392–403.

——, P. J. Lenk 1994. Modeling household purchase behavior with logistic normal regression. *J. American Statistical Association.* 89 December 1218–1231.

Amemiya, T. 1981. Qualitative response models: A survey. *J. Economic Literature.* 19 December 1483–1536.

Aptech 1995. *Gauss Applications.* Aptech Systems, Inc., Maple Valley, WA.

Bekker, P. A., A. Merckens, T. J. Wansbeek 1994. *Identification, Equivalent Models, and Computer Algebra.* Academic Press, San Diego, CA.

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

Börsch-Supan, A., V. A. Hajivassiliou 1993. Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. *J. Econometrics.* 58 3 347–368.

——, ——, L. J. Kotlikoff, J. N. Morris 1990. Health, children, and elderly living arrangements: A multiperiod, multinomial probit model with unobserved heterogeneity and autocorrelated errors, NBER, Working paper 3343, Cambridge, MA.

Bunch, D. S. 1991. Estimability in the multinomial probit model. *Transportation Res. B.* 25B 1 1–12.

——, R. Kitamura 1991. 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., P. E. Green 1995. Psychometric methods in marketing research: Part 1, conjoint analysis. *J. Marketing Res.* 32 November 385–391.

Chintagunta, P. K. 1992. Estimating a multinomial probit model of brand choice using the method of simulated moments. *Marketing Sci.* 11 4 386–407.

——, B. E. Honoré 1996. Investigating the effects of marketing variables and unobserved heterogeneity in a multinomial probit model. *International J. Res. in Marketing.* 13 1 1–15.

Cohen, S. H. 1997. Perfect union: CBCA marries the best of conjoint and discrete choice models. *Marketing Res.* 9 1 12–17.

Currim, I. S. 1982. Predictive testing of consumer choice models not subject to independence of irrelevant alternatives. *J. Marketing Res.* 19 2 208–222.

Daganzo, C. D. 1979. *Multinomial Probit, The Theory and Its Applications to Demand Forecasting.* Academic Press, New York.

DeSarbo, W. S., V. Ramaswamy, S. H. Cohen 1995. Market segmentation with choice-based conjoint analysis. *Marketing Letters.* 6 2 137–148.

Elrod, T., M. P. Keane 1995. A factor-analytic probit model for representing the market structures in panel data. *J. Marketing Res.* 32 1 1–16.

——, J. J. Louviere, K. S. Davey 1992. An empirical comparison of rating-based and choice-based conjoint models. *J. Marketing Res.* 29 3 368–377.

Geweke, J., M. P. Keane, D. Runkle 1994. Alternative computational approaches to inference in the multinomial probit model. *Rev. Economics and Statistics.* 76 4 609–632.

Gönül, F., K. Srinivasan 1993. Modeling multiple sources of heterogeneity in multinomial logit models: Methodological and managerial issues. *Marketing Sci.* 12 3 213–229.

Green, P. E., A. M. Krieger 1993. Conjoint analysis with product-positioning applications. J. Eliashberg, G. L. Lilien, eds. *Handbooks in Operations Research and Management Science,* Vol. 5, Marketing.

——, V. Srinivasan 1978. Conjoint analysis in consumer research: Issues and outlook. *J. Consumer Res.* 5 September 103–123.

Hajivassiliou, V. A. 1993. Simulation estimation methods for limited dependent variable models. G. S. Maddala, C. R. Rao, H. D. Vinod, eds. *Handbook of Statistics*, Vol. 11, 519–543.

——, D. McFadden 1990. The method of simulated scores for the estimation of LDV models. Working paper, Yale University, New Haven, CT.

——, ——, P. Ruud 1993. Simulation of multivariate normal rectangle probabilities and their derivatives: Theoretical and computational results. Working paper, Yale University, New Haven, CT.

Hausman, J. A., D. A. Wise 1978. A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. *Econometrica.* 46 2 403–426.

Heckman, J. J., G. Sedlacek 1985. Heterogeneity, aggregation, and market wage function: An empirical model of self-selection in the labor market. *J. Political Economy.* 93 6 1077–1125.

Huber, J., J. W. Payne, C. Puto 1982. Adding asymmetrically dominant alternatives: Violations of regularity and the similarity hypotheses. *J. Consumer Res.* 9 90–98.

——, C. Puto 1983. Market boundaries and product choice: Illustrating attraction and substitution effects. *J. Consumer Res.* 10 31–44.

Kamakura, W. A. 1989. The estimation of multinomial probit models: A new calibration algorithm. *Transportation Sci.* 23 4 253–265.

——, R. K. Srivastava 1984. Predicting choice shares under conditions of brand interdependence. *J. Marketing Res.* 21 November 420–434.

Keane, M. P. 1992. A note on identification in the multinomial probit model. *J. Business and Economic Statistics.* 10 2 193–200.

——, 1997. Current issues in discrete choice modeling. *Marketing Letters.* 8 3 307–322.

Lee, L.-F. 1992. On efficiency of methods of simulated moments and maximum simulated likelihood estimation of discrete response models. *Econometric Theory.* 8 518–552.

Lenk, P. J., W. S. DeSarbo, P. E. Green, M. R. Young 1996. Hierarchical bayes conjoint analysis: Recovery of partworth heterogeneity from reduced experimental designs. *Marketing Sci.* 15 2 173–191.

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

——, G. Woodworth 1983. Design and analysis of simulated consumer choice or allocation experiments: An approach based on aggregate data. *J. Marketing Res.* 20 4 350–367.

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

Mahajan, V., P. E. Green, S. M. Goldberg 1982. A conjoint model for measuring self and cross-price/demand relationships. *J. Marketing Res.* 19 3 334–342.

McCulloch, R., P. E. Rossi 1994. An exact likelihood analysis of the multinomial probit model. *J. Econometrics.* 64 207–240.

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

—— 1989. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica.* 57 5 995–1026.

Mühleisen, M. 1991. On the use of simulated estimators for panel models with limited dependent variables. Working paper, University of Munich, Germany.

Nowlis, S. M., I. Simonson 1997. Attribute-task compatibility as a determinant of consumer preference reversals. *J. Marketing Res.* 34 May 205–218.

Papatla, P. 1996. A multiplicative fixed-effects model of consumer choice. *Marketing Sci.* 15 3 243–261.

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

Rossi, P. E., G. M. Allenby 1993. A Bayesian approach to estimating household parameters. *J. Marketing Res.* 30 171–182.

——, R. E. McCulloch, G. M. Allenby 1996. The value of purchase history data in target marketing. *Marketing Sci.* 15 4 321–340.

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. *J. Consumer Res.* 16 158–174.

——, A. Tversky 1992. Choice in context: Tradeoff contrasts and extremeness aversion. *J. Marketing Res.* 29 281–295.

Tversky, A. 1972. Elimination by aspects: A theory of choice. *Psychological Rev.* 79 281–299.

Vriens, M., M. Wedel, T. Wilms 1996. Metric conjoint segmentation methods: A Monte Carlo comparison. *J. Marketing Res.* 32 1 73–85.

This paper was received February 13, 1997, and has been with the authors 12 months for 2 revisions; processed by Greg M. Allenby.