Intern. J. of Research in Marketing 25 (2008) 94–103

Contents lists available at ScienceDirect

Intern. J. of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Models and optimal designs for conjoint choice experiments including a no-choice option

Bart Vermeulen [a,*,1], Peter Goos [b], Martina Vandebroek [c]

[a] Faculty of Business and Economics, Katholieke Universiteit Leuven, Belgium
[b] Faculty of Applied Economics, Universiteit Antwerpen, Belgium
[c] Faculty of Business and Economics, University Center for Statistics, Katholieke Universiteit Leuven, Belgium

ARTICLE INFO

Article history:
First received in 19 January 2007
and was under review for 3 months

Area Editor: Peter Danaher

Keywords:
Bayesian optimal design
Choice-based conjoint
Conjoint analysis
D-optimality
Model-robust design
Multinomial logit model
Nested logit

ABSTRACT

In a classical conjoint choice experiment, respondents choose one profile from each choice set that has to be evaluated. However, in real life, the respondent does not always make a choice: often he/she does not prefer any of the options offered. Therefore, including a no-choice option in a choice set makes a conjoint choice experiment more realistic. In the literature, three different models are used to analyze the results of a conjoint choice experiment with a no-choice option: the no-choice multinomial logit model, the extended no-choice multinomial logit model, and the nested no-choice multinomial logit model. We develop optimal designs for the two most appealing of these models using the *D*-optimality criterion and the modified Fedorov algorithm and compare these optimal designs with a reference design, which is constructed while ignoring the no-choice option, in terms of estimation and prediction accuracy. We conclude that taking into account the no-choice option when designing a no-choice experiment only has a marginal effect on the estimation and prediction accuracy as long as the model used for estimation matches the data-generating model.

1. Introduction

The aim of a conjoint choice experiment is to model respondents' choices as a function of the features of a product or service. In this type of experiment, each respondent repeatedly chooses the option offering the maximum amount of utility from each of a number of choice sets, each containing several options. In past decades, these experiments have become increasingly popular for modelling market demand (see e.g. Kamakura, Wedel, & Agrawal, 1994; Wedel, Vriens, Bijmolt, Krijnen, & Leeflang, 1998) because of their ability to simulate market decisions realistically and because of the opportunity to estimate the impact of product or service features on market demand.

In a classical conjoint choice experiment, the respondent is forced to choose one profile from each choice set. However, in real life the customer does not always make a choice: often he/she does not like any of the options presented and does not buy any of the products or services offered. Therefore, including a no-choice option in a choice set makes the experiment more realistic.

To conduct an efficient conjoint choice experiment with a small number of choice sets, an optimal design has to be developed by choosing the appropriate alternatives and grouping them in choice sets in the best possible way. We examine whether an optimal no-choice design, i.e. a design constructed taking into account the presence of a no-choice option in the choice sets of the experiment, leads to better results in terms of the accuracy of the estimated model coefficients and the predicted probabilities in comparison to a reference design developed ignoring the no-choice option.

In the next section we discuss the respondents' motivation to choose the no-choice option and the advantage and disadvantage of including a no-choice option in a choice set. In Section 3, three models for analyzing the data from a conjoint choice experiment with a no-choice option are discussed: the no-choice multinomial logit model (NCMNL), the extended no-choice multinomial logit model (ENCMNL), and the nested no-choice multinomial logit model (NLMNL). In Section 4, we explain some basic notions of experimental design and introduce the *D*-optimality criterion, which we apply to the ENCMNL and NLMNL models to develop optimal no-choice designs. Furthermore, we create a no-choice design which is robust against the data-generating model—i.e. the behavioral model driving the choices of the respondents, and a reference design. The relative performances of the reference design and the no-choice designs under different scenarios are compared in Section 5. In Section 6, we use a simulation study to measure the accuracy of the parameter estimates by the expected mean squared error of the parameter estimates and the prediction accuracy of the designs by the expected mean squared error of the predicted probabilities. Finally, in Section 7 we take a detailed look at the

* Corresponding author. Department of Decision Sciences and Information Management, Faculty of Business and Economics, Katholieke Universiteit Leuven, Naamsestraat 69, 3000 Leuven, Belgium. Tel.: +32 16 32 69 63, fax: +32 16 32 66 24.
E-mail address: bart.vermeulen@econ.kuleuven.be (B. Vermeulen).
[1] Bart Vermeulen is funded by project G.0611.05 of the Fund for Scientific Research Flanders.

0167-8116/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijresmar.2007.12.004

accuracy of the predictions based on a simulation study in which the data consists of a mixture of choices generated by the ENCMNL and NLMNL models.

2. The no-choice option

In this section, we discuss several aspects of the no-choice option described in the literature. First, we focus on the reasons why this option is attractive to respondents. Subsequently, we discuss the advantage and disadvantage of incorporating this option in the design and the model.

In the literature that deals with the no-choice option in choice experiments, two reasons why a respondent would choose the no-choice option can be found. According to the rationale theory, which reduces decision-making to the concept of utility, the consumer prefers the product that offers him/her the maximum amount of utility. None of the alternatives is considered attractive when none of them offers the respondent sufficient utility. In that case, the benefits of continuing the respondent's search for better alternatives are greater than the costs. For this reason, the respondent chooses the no-choice option and looks for more useful alternatives. Psychological research provides another theory as to why a consumer chooses not to choose. The theory focuses strongly on avoiding intricate trade-offs and the related discomfort and fear of making the wrong choice. Baron & Ritov (1994), for example, state that consumers prefer bearing the consequences of inaction rather than those of wrong action. This is the reason why a consumer prefers deferring his purchase over buying the wrong product or service when he/she feels uncomfortable choosing. In this situation, the no-choice option is used as a way to avoid a choice conflict between two alternatives with nearly equal utilities. Johnson & Orme (1996), however, found no evidence for such behavior and claim that respondents tend not to choose the no-choice option to avoid difficult decision-making.

In this article, we therefore assume that the respondents determine the utility for each option and choose the no-choice option if none of the alternatives offers sufficient utility. Consequently, the meaning of the no-choice option given in this paper is "None of the alternatives meet my requirements," signifying that the customer prefers to continue to look for better alternatives. The rationale theory enables us to use the multinomial logit model which is the focus of the following sections (Dhar, 1997; Dhar & Simonson, 2003).

The major advantage of including a no-choice option in a conjoint choice experiment is that a more realistic experiment is obtained. The experiment, therefore, leads to better estimates of the model parameters and to better predictions of market penetrations. As a matter of fact, forcing a respondent to make a choice in a conjoint choice experiment might lead to biased parameters when analyzing the choice data (Dhar, 1997; Dhar & Simonson, 2003). That including a no-choice option in the experiment avoids the bias is a major advantage, which should outweigh the disadvantage that, each time a respondent selects the no-choice option, no information is collected concerning the relative attractiveness of the alternatives offered.

3. Multinomial logit models

In this section, we discuss the multinomial logit model and the nested multinomial logit model. For each of these models, we review the logit probability of choosing an alternative and the likelihood function of the corresponding model. Within the class of multinomial logit models, Haaijer, Kamakura, and Wedel (2001) describe two models for analyzing data from choice experiments that have a no-choice option: the NCMNL model and ENCMNL model. The use of these two models, described in Section 3.1, requires the "independence of irrelevant alternatives" assumption to be valid. The violation of this assumption necessitates the use of the nested logit model, which is the subject of Section 3.2. In this paper, we refer to the nested logit model as the NLMNL model.

3.1. The NCMNL and ENCMNL models

The most popular model to analyze choice data is the multinomial logit model. If a respondent n faces choice set k with J alternatives, then the utility of the jth alternative of that choice set experienced by respondent n can be expressed as

$$u_{nkj} = \boldsymbol{x}'_{kj}\boldsymbol{\beta} + \varepsilon_{nkj}. \tag{1}$$

The p-dimensional parameter vector $\boldsymbol{\beta}$, the elements of which are often referred to as part-worths, contains the importance of the attributes for the consumer in determining his/her utility. We assume that this vector is common for all respondents. The vector $\boldsymbol{x}_{kj}$ has the same dimension as $\boldsymbol{\beta}$ and contains the levels of the attributes of the jth alternative in choice set k. The error term ε_{nkj} captures the influence of unobserved factors on the utility experienced by the respondent. All error terms are assumed to be independent and identically extreme-value distributed. Under this assumption, the probability that respondent n chooses alternative j of choice set k is

$$P_{nkj} = \frac{\exp\left(\boldsymbol{x}'_{kj}\boldsymbol{\beta}\right)}{\sum_{i=1}^{J}\exp\left(\boldsymbol{x}'_{ki}\boldsymbol{\beta}\right)}. \tag{2}$$

If we assume that N respondents evaluate the same set of K choice sets, the log-likelihood function for the multinomial logit model becomes

$$\ln\ (L(\boldsymbol{\beta})) = \sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{j=1}^{J} y_{nkj}\ \ln\ \left(P_{nkj}\right). \tag{3}$$

The dummy variable y_{nkj} equals one when respondent n prefers alternative j of choice set k and zero otherwise. The maximum likelihood estimate $\hat{\boldsymbol{\beta}}$ for the parameter vector is obtained by maximizing the log-likelihood function.

Like Haaijer et al. (2001), we distinguish two multinomial logit models for analyzing data from choice experiments with a no-choice option. The simplest model, the NCMNL model, represents the no-choice option by an alternative having zero values for all attribute levels. Consequently, the deterministic part of the utility of the no-choice option in the NCMNL model is always zero. This method of coding the no-choice option possibly leads to distorted parameter estimates when linear attributes are present.

In the second model, the ENCMNL model, an extra no-choice dummy variable is used to represent the no-choice option. This offers the advantage of an enhanced model fit. The no-choice dummy variable acts as an additional two-level attribute and takes value zero for all real-choice options and value one for the no-choice option. The extra model parameter corresponding to the dummy variable is interpreted as the utility of choosing the no-choice option by the respondent. Contrary to the NCMNL model, the deterministic part of this utility is not equal to zero and consequently, the probability of choosing this option differs between these two models. As the primary interest of researchers is not in the estimation of the extra model parameter, we develop optimal designs that focus on the precise estimation of the other model parameters, i.e. the part-worths of the original attributes. As the NCMNL model is nested within the ENCMNL model and leads to a poorer model fit, we do not consider the NCMNL model in detail in the following sections.

A problem with both these models is that they require the strong assumption of the independence of irrelevant alternatives, commonly referred to as the IIA-assumption, to be valid. Under this assumption, the relative odds of choosing alternative j over j' depend only on the attributes of j and j' and not on the attributes of the other alternatives. This implies that the unobserved parts of the utilities of the alternatives exhibit no correlation. While the IIA-assumption is a realistic one in