Food Quality and Preference 54 (2016) 79–89

Contents lists available at ScienceDirect

Food Quality and Preference

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

Comparison of two different strategies for investigating individual differences among consumers in choice experiments. A case study based on preferences for iced coffee in Norway

D. Asioli [a,c,*], V.L. Almli [a], T. Næs [a,b]

[a] Consumer and Sensory Science – Division Food Science, Nofima AS, PO Box 210, 1431 Ås, Norway
[b] University of Copenhagen, Faculty of Life Sciences, Department Food Science, Rolighedsvej 30, 1958 Fredriksberg Copenhagen, Denmark
[c] School of Economics and Business, Norwegian University of Life Sciences, Universitetstunet 3, Ås, Norway

ARTICLE INFO

Article history:
Received 9 May 2016
Received in revised form 5 July 2016
Accepted 6 July 2016
Available online 9 July 2016

Keywords:
Consumer attributes
Choice experiment
Method comparison
Mixed Logit Model
Iced coffee

ABSTRACT

Two different strategies for investigating individual differences among consumers in choice experiments using the Mixed Logit Model are compared. The study is based on a consumer study of iced coffees in Norway. Consumers (n = 102) performed a choice task of twenty different iced coffee profiles varying in coffee type, production origin, calorie content and price following an orthogonal design. Consumer attributes, such as socio-demographics, attitudes and habits, were also collected. Choice data were first analyzed using the Mixed Logit Model and then two different approaches were adopted for investigating consumer attributes. The first strategy, called *one-step strategy*, includes the consumer attributes directly in the Mixed Logit Model. The second strategy, called *multi-step strategy*, combines different methods of analysis such as Mixed Logit Model based on the design factors only, followed by Principal Component Analysis and Partial Least Squares regression to study consumer attributes. The two approaches are compared in terms of data analysis methodologies, outcomes, practical issues, user friendliness, and interpretation. Overall, we think the *multi-step* strategy is the one to be preferred in most practical applications because of its flexibility and stronger exploratory capabilities.

1. Introduction

1.1. Conjoint analysis (CA)

One of the most frequently used methodologies for consumer studies is conjoint analysis (CA). This is a method which is able to estimate the structure of consumer evaluations using a set of product profiles consisting of predetermined combinations of product attributes (Green & Srinivasan, 1990). Consumers are presented with these product profiles and are asked to either rank, rate or choose among them (Louviere, Hensher, & Swait, 2000; Molteni & Troilo, 2007). Within CA there are two main categories: (i) acceptance-based approaches, which require that consumers rate each alternative product according to their degree of liking or hypothetical purchase intention and (ii) preference-based approaches, where consumers are required to express their preferences either in terms of ranks or of choices among several alternative products with varying levels of attributes. In this paper we will focus on the choice approach.

1.2. Choice experiment (CE)

Choice based experiments (CEs) have been developed for investigating consumers' choice both for market and non-market goods (Haaijer, Kamakura, & Wedel, 2001; Louviere et al., 2000; Yangui, Akaichi, Costa-Font, & Gil, 2014). In a choice study, consumers are presented with a series of alternative choice scenarios and are asked to choose their most preferred option within each choice scenario. The different alternatives are composed of different combinations of attribute levels which characterize the goods (e.g. price, nutritional content, etc.) usually based on an experimental design. One of the arguments put forward for choice-based methods in comparison to rating or ranking methods, is that having respondents choose a single preferred stimulus among a set of stimuli better approximates a real purchase situation (Carson et al., 1994; Louviere et al., 2000). CEs originate from economics and are increasingly expanding to different fields such as transportation, environment, health and marketing. During the last

* Corresponding author at: Consumer and Sensory Science – Division Food Science, Nofima AS, PO Box 210, 1431 Ås, Norway.
E-mail address: daniele.asioli@nofima.no (D. Asioli).

http://dx.doi.org/10.1016/j.foodqual.2016.07.005
0950-3293/

years there have been an increasing number of applications of CEs also in food consumer studies (Lusk, Fields, & Prevatt, 2008; Van Loo, Caputo, Nayga, Meullenet, & Ricke, 2011; Van Wezemael, Caputo, Nayga, Chryssochoidis, & Verbeke, 2014).

1.3. Consumer heterogeneity

Consumer heterogeneity with respect to preference pattern, described as "a key and permanent feature of food choice" by Combris, Bazoche, Giraud-Héraud, and Issanchou (2009), is an important and natural element of food choice research (Almli, Øvrum, Hersleth, Almøy, & Næs, 2015). Preference heterogeneity can be investigated in terms of demographics (e.g. gender, age, income), attitudes (e.g. preference for certain product characteristics) and habits (e.g. ways and location of food consumption), and is of particular importance for food practitioners (Næs, Brockhoff, & Tomic, 2010) in order to develop and market food products that better meet consumers' needs and wishes.

At an overall level and independently from data collection and statistical approach, one can identify two main strategies of consumers segmentations: *a priori* segmentation and *a posteriori* segmentation (Næs, Kubberød, & Sivertsen, 2001; Næs et al., 2010). The *a priori* segmentation is based on splitting the consumer group into segments according to consumer attributes and then analyzing the group preferences separately or together in an ANOVA model or a Mixed Logit Model (depending on data collection, see e.g. Asioli, Næs, Øvrum, & Almli, 2016) that combine design factors and consumer attributes in one single model (Næs et al., 2010).

The second strategy is called *a posteriori* segmentation and is based on creating consumer groups of similar product preferences by analyzing the actual preference, liking or purchase intent data to create segments, and then relating segments to consumer characteristics *a posteriori*. According to Gustafsson, Herrmann, and Huber (2003) there are different approaches to *a posteriori* segmentation. The main advantage of *a posteriori* segmentation is that it is unsupervised in the sense that the segments are determined without external influence of consumer attributes, so it is more open to new and unexpected results (Næs et al., 2010). In this paper we will use an approach based on visual inspection of scores plots from principal components analysis (PCA) (see e.g. Endrizzi, Gasperi, Rødbotten, & Næs, 2014), but other possibilities also exist. An important example here is Latent Class Analysis (LCA) which is based on a mathematical optimisiation criterion developed for splitting the group of consumers into segments with similar response pattern (Boxall & Adamowicz, 2002).

It should be mentioned that there also exists another option more or less between the two segmentation strategies discussed above. This is based on using the consumer attributes explicitly in the segmentation procedure as done in for instance by Vigneau, Endrizzi, and Qannari (2011). In this paper, however, only *a priori* and *a posteriori* segmentation will be in focus.

1.4. Objectives of the study

The objective of this study is to compare two different strategies of investigating consumer attributes in CEs, one *a priori* and one *a posteriori* strategy. The first strategy includes consumer attributes *a priori* together with product attributes in a Mixed Logit Model and is therefore a one-step strategy. The second strategy is a two-step strategy based on investigating consumers with similar/dissimilar choices using a Mixed Logit Model followed by Principal Component Analysis (PCA) and partial least squares (PLS) regression (Wold, Martens, & Wold, 1983) or PLS classification (Ståhle & Wold, 1987) for relating the preference pattern to the consumer attributes *a posteriori*. To compare the methods, data from a conjoint choice experiment investigating consumer preferences for iced coffee products in Norway were used. Practical issues, user-friendliness and interpretation of the two approaches will be discussed.

2. Theory: Statistical methods used

Choice-based data are routinely analyzed within a random utility framework called Discrete Choice Models (DCMs) (Train, 2009). The approach is based on modelling "utility", that is to say the net benefit a consumer obtains from selecting a specific product in a choice situation, as a function of the conjoint factors. DCMs aim at understanding the behavioural process that leads to a consumer's choice (Train, 2009). DCMs emerged some decades ago and have undergone a rapid development from the original fixed coefficients models such as multinomial logit, to the highly general and flexible Mixed Logit (ML) model. In the ML model, the utility of a product *j* for individual *m* in a choice occasion *t* is written:

$$U_{mjt} = \beta'_m x_{mjt} + \varepsilon_{mjt} \tag{1}$$

where β_m is a random vector of individual-specific parameters accounting for preference heterogeneity, x_{mjt} is a vector of conjoint factors, and ε_{mjt} is a random error term. For the ML model it is assumed that the random errors are independent identically distributed (i.i.d) and follow a so-called extreme value distribution (see Train, 2009 for theoretical argument for the distributional assumption). An advantage of the ML model is that one may freely include random parameters β_m of any distributions and correlations between random factors. This flexibility allows writing models that better match real-world situations. ML models have been applied also in consumer food studies (Alfnes, 2004; Bonnet & Simioni, 2001; Hasselbach & Roosen, 2015; Øvrum, Alfnes, Almli, & Rickertsen, 2012). In Øvrum et al. (2012) CE was used for investigating how diet choices are affected by exposure to diet-related health information on semi-hard cheese. Hasselbach and Roosen (2015) investigated whether the concepts of organic and local food support or threaten each other in consumers' choice by using a CE. Alfnes (2004) investigated Norwegians consumers' preferences for country of origin and hormone status of beef using the ML model. In these studies, as in most studies which apply the ML model, consumers' heterogeneity was not investigated in depth (i.e. segmentation).

In the next two sections (2.1 and 2.2), the two strategies introduced in Section 1.3 will be described.

2.1. Strategy 1: Simultaneous Mixed Logit Model of the conjoint factors and consumer attributes (One-step strategy with a priori segmentation)

The first strategy is inspired by the analysis of individual acceptance ratings using a Mixed Model ANOVA approach (see e.g. Næs, Almli, Bølling Johansen, & Hersleth, 2010). It consists of including both conjoint factors and categorical consumer characteristics and their interactions in one model. This means that in addition to the conjoint factor x_{mjt} in the model above, one adds additional variables that represent the consumer attributes. In practice, the number of attributes added in this way should be limited due to the lowering of power and also possible more complex interpretation. Note that attributes added in this way could also in principle be based on consumer segments (obtained by for instance an initial analysis) other than those obtained by using the measured consumer attributes individually.

Note that interactions between conjoint factors and consumer attributes are of special importance since they represent how the different consumer groups respond differently to the different