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

Article

Accepted Version

Asioli, D., Almli, V. L. and Næs, T. (2016) 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. Food Quality and Preference, 54. 79 - 89. ISSN 0950-3293 doi: https://doi.org/10.1016/j.foodqual.2016.07.005 Available at http://centaur.reading.ac.uk/75176/

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

Asioli, D. [1,3], Almli, V.L. [1], Næs, T. [1,2]

[1]Consumer and Sensory Science - Division Food Science

Nofima AS, PO Box 210, 1431 Ås, Norway

[2]University of Copenhagen, Faculty of Life Sciences,

Department Food Science, Rolighedsvej 30, 1958 Fredriksberg Copenhagen, Denmark

[3]School of Economics and Business, Norwegian University of Life Sciences,

Universitetstunet 3, Ås, Norway

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 analysed 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, Hensher, & Swait, 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 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, & 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, et al., 2010; Næs, Kubberød, & Sivertsen, 2001). 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, A., Herrmann, A., & 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, & 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 analysed 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:

$$U_{mjt} = \beta'_m x_{mjt} + \varepsilon_{mjt} \qquad (1)$$

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 & 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 $\mathbf{x}_{\mathbf{mjt}}$ 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 conjoint factors. This strategy is the same as used in Asioli et al. (2016) for analsying the same data set as used here.

2.2 STRATEGY 2: Combining Mixed Logit model, PCA and PLS regression (Multi-step strategy with a posteriori segmentation)

The second strategy has been initially proposed within the framework of Mixed Model ANOVA (Endrizzi et al., 2014; Endrizzi, Menichelli, Johansen, Olsen, & Næs, 2011; Næs, Almli, et al., 2010). However, this approach can also easily be extended to choice data using the Mixed Logit model (Almli et al., 2015). First, choice data are analyzed using the ML model by including only conjoint factors and possibly also their interactions, as presented in Eq. 1). Then, the matrix of individual parameter estimates $\hat{\boldsymbol{\beta}}_m$ extracted from the ML model are analyzed and interpreted using Principal Component Analysis (PCA). At this point, two different approaches for investigating consumer attributes can be applied.

Option 1. A first possible approach is to relate the PCs directly to consumer attributes using for instance Partial Least Squares regression (Endrizzi et al., 2011) which can easily handle a large number of highly collinear attributes. Note that one could also use the parameter estimates $\hat{\boldsymbol{\beta}}_m$ directly as responses in the PLS regression or several principal components at the same time. The choice made here of using the PCs as dependent variables was made since the principal components correspond more or less 100% to the design variables, and since it is of major interest to investigate explicitly how the consumer attributes relate to the different conjoint factors in the design. This option also facilitates the comparison with the first analysis strategy described above (Strategy 1). In order to highlight this aspect, each principal component was handled independently.

Option 2. A second possible approach is to identify segments in the Principal Component Analysis (PCA), either visually (visual segmentation, Endrizzi et al., 2011) or automatically (using cluster analysis). Then, the consumer segments are investigated in terms of socio-demographics, habits and attitudes attributes using for instance Partial Least Squares – Discrimintion Analysis (PLS-DA, Barker & Rayens, 2003; Ståhle & Wold, 1987) which relates the consumer segments to consumer attributes. The main advantage of such an approach is that one can decide during the second step which segments or groups of consumers one is interested in investigating. An application of this method is provided by Almli et al. (2015) who used this approach on ranking data in a consumer study of semi-hard cheese.

In this paper, all PLS regressions and PLS-DA models were run on standardised input variables, using cross-validation on 10 random segments and performing a jack-knife uncertainty test with 95% confidence interval for the detection of significant variables (Martens & Martens, 2000). Calculations were performed in The Unscrambler X 10.2 (Camo Software AS, Oslo). Due to the large number of consumer attributes collected, a two-step procedure was used: in the first step all the consumers' attributes were included in the model. Then, in the second step a new model was run only including significant consumers'attributes from the first step. This results in a better suited and more parsimonial model. For the PLS-DA the consumer groups were represented by dummy variables (Ys) in the PLS-DA, while consumer attributes were used as independent variables (Xs).

3. MATERIAL AND METHODS

3.1 Consumer test

We tested the approaches using a dataset based on iced-coffee products. A sample of 102 consumers was recruited in the region south of Oslo, Norway, in November 2012. The test included four sessions, one of them being a choice task. For details about the experiment and socio – demographic characteristics of the sample investigated, see Asioli et al. (2016).

3.2 Iced coffee products

Conjont factors and their levels for the iced coffee profiles presented to the consumers were selected based on focus group results; see Asioli, Næs, Granli, & Lengard Almli (2014) for details. Table 1 shows the four conjont factors and levels that were selected: coffee type, calorie content and origin with two levels each, and price with three levels.

Table 1 – Conjont factors and levels used in the conjoint design

<<Please, place here table 1>>

3.3 Choice task

An orthogonal choice design composed of eight choice sets of three products each was generated in SAS version 9.3 (see appendix I). The design featured 20 unique samples where all of them were taken from the full factorial design (see Asioli et al, 2016 for more details) . Usually in choice studies a “no-choice” option is included because it can provide a better market penetration prediction (Enneking et al., 2007; Haaijer et al., 2001). However, in this paper we did not aim to predict market penetration, thus we decided not to include the “no-choice” option and only iced-coffee consumers were recruited to the test.

The eight triads of iced coffee profiles were displayed successively on a computer screen in the form of photographs (see Figure 1).

Figure 1 – One of the iced coffee profiles

<<Please, place here figure 1>>

Product presentation was randomized across participants both at choice set level and at product level within choice sets. For each choice-set, consumers' probability of buying was elicited with the question: *"Imagine that you are purchasing iced coffee. Which of these iced coffees are you most likely to buy?"* and participants answered by clicking on one of the three alternatives.

3.4 Consumer attributes

In order to investigate individual differences, we have collected a number of consumer attributes. The attributes investigated are related to iced coffee consumption habits (importance of attributes for purchasing, consumption frequency, duration (years) of iced coffee consumption, consumption time of the day, location of consumption, location of purchasing, alternative products, motivations of consumption and types of products), warm coffee habits (types of additives, location of consumption), food attitudes (items of food neophobia, health consciousness and ethnocentricity) and socio-demographic attributes.

Consumers attributes are measured using both numerical and categorical variables. For the importance of attributes for choosing iced coffee, the scale is anchored in 1 (Not important at all) and 5 (Very important at all). The same is the case for the habits attributes. All the categorical attributes have been coded using dummy variables where 0 represents the absence of the actual level while 1 represents the presence of the attribute level. The complete list of attributes can be obtained from the authors.

3.5 Data analysis

All conjoint factors were coded using effects coding (-1; 1) (Bech & Gyrd-Hansen, 2005), except price which was coded in three levels (mean centered) (-1; 0; 1). In other words, the price was coded as a linear covariate (see Asioli et al., 2016 for arguments). For illustration of Strategy 1, we decided to consider only two segmentation attributes, Gender and Age group. Note that many other choices could have been made, these two are only chosen for illustration of the methodology. The factors used were coded as presented in Table 2.

Table 2 – Factors coded and their description

<<Please, place here table 2>>

The ML model for the two cases considered here provide both population averages of the regression coefficients and the set of individual coefficients. The population averages can be interpreted directly in terms of p-values and their signs. Magnitudes of the factors can only be considered relative to one another since the utility scale does not represent a true rating scale given by the consumers (see Train, 2009). The standard deviation of the individual coefficients will also be considered in this paper.

3.5.1 STRATEGY 1: Simultaneous Mixed Logit model of the conjoint factors and consumer attributes (One-step strategy)

Following eq. 2) below, we included two consumer attributes in the ML model, namely Gender and Age. Introducing more consumer attributes may make the estimated conjoint effects weaker and thus disturb interpretation (Næs, Almli, et al., 2010); it may also be technically more difficult to achieve in a software context. This is particularly true if there are attributes with several levels or attributes that are continuous. In addition, the attributes may be collinear, making estimation very unstable and the results difficult to interpret. In this paper we confine ourselves to incorporating two consumer attributes Gender and Age.

In our main specification of the model we incoporate main effects of the conjoint factors and all two-factor interactions among the conjoint factors and between the conjoint factors and the consumer attributes. The utility ML model for iced coffee *j* for individual *i* in choice occasion *t* can be written:

$$\begin{aligned} U_{ijt} = {} & \beta_{1i}\,Coffee_{ijt} + \beta_{2i}\,Calories_{ijt} + \beta_{3i}\,Origin_{ijt} + \beta_{4i}\,Price_{ijt} + \beta_{5i}\,(Coffee{*}\,Calories)_{ijt} + \beta_{6i} \\ & (Coffee{*}Origin)_{ijt} + \beta_{7i}\,(Coffee{*}Price)_{ijt} + \beta_{8i}\,(Calories{*}Origin)_{ijt} + \beta_{9i} \\ & (Calories{*}Price)_{ijt} + \beta_{10i}\,(Origin{*}Price)_{ijt} + \beta_{11i}\,(Age{*}Coffee)_{ijt} + \beta_{12i}\,(Age\,{*}Price)_{ijt} + \\ & \beta_{13i}\,(Age\,{*}Calories)_{ijt} + \beta_{14i}\,(Age{*}Origin)_{ijt} + \beta_{15i}\,(Gender{*}Coffee)_{ijt} + \beta_{16i} \\ & (Gender{*}Price)_{ijt} + \beta_{17i}\,(Gender\,{*}Calories)_{ijt} + \beta_{18}\,(Gender{*}Origin)_{ijt} + \varepsilon_{mjt} \end{aligned} \qquad (2)$$

The interaction effects are obtained by multiplying the columns in the data set for the corresponding main effects. The consumer effect is automatically incorporated here since all coefficients are considered random. Note that Gender and Age have no main effect, the reason being that only the relative differences in each individual's utility pattern influences the choice model. The chosen ML model assumes independent random parameters with normal distributions for all conjoint factors, consumer attributes and two-way interactions. The ML model was estimated using the Stata module *mixlogit* (Hole, 2007) run in STATA 11.2 software (StataCorp LP, College Station, US). Four thousand Halton draws were used in the simulations. More details on estimation of ML models are found in Train (2009) and Hole (2007). Note that from a segments point of view the interest lies in the interactions between consumer attributes and the conjoint factors. Note also that one can calculate the individual random coefficients and their standard deviations (SDs) for this model as will be shown in Section 4.1.

3.5.2 STRATEGY 2: Mixed Logit Model, PCA and PLS (Multi-step strategy)

Mixed Logit Model

Following eq.1), we developed a Mixed Logit Model which includes the main effects and two-way interactions among conjoint factors. Thus, in our main specification of the model we included all the main effects and interactions among the conjoint factors for Coffee, Calories, Origin and Price. The utility ML model for iced coffee *j* for individual *i* in choice occasion *t* is written:

$$U_{ijt} = \beta_{1i}\, Coffee_{ijt} + \beta_{2i}\, Calories_{ijt} + \beta_{3i}\, Origin_{ijt} + \beta_{4i}\, Price_{ijt} + B_{5i}\, (Coffee*Calories)_{ijt} + \beta_{6i}\, (Coffee*Origin)_{ijt} + \beta_{7i}\, (Coffee*Price)_{ijt} + \beta_{8i}\, (Calories*Origin)_{ijt} + \beta_{9i}\, (Calories*Price)_{ijt} + \beta_{10i}\, (Origin*Price)_{ijt} + \varepsilon_{mjt} \quad (3)$$

As can be seen, except for the consumer attributes, the two models are identical. For the technical details on how the calculations have been performed see section 3.5.1.

Then, the matrix of individual parameter estimates $\hat{\boldsymbol{\beta}}_m$ was extracted from the ML model (Eq. 3) by using the command *mixlbeta* in STATA. Note that this matrix of individual estimates plays a similar role as the residuals matrix from a reduced mixed model ANOVA on rating data in the sense that both reflect individual variations from population effects (Næs, Almli, et al., 2010).

Principal Component Analysis (PCA)

The matrix of individual parameter estimates $\hat{\boldsymbol{\beta}}_m$ extracted from the Mixed Logit Model analysis is submitted to Principal Component Analysis (PCA) in order to identify the main components of variation between individuals. PCA was conducted in the multivariate statistical software package The Unscrambler X 10.2 (Camo Software AS, Norway).

Partial Least Squares (PLS) regression

PLS regression was conducted in the multivariate statistics software package The Unscrambler X 10.2 (Camo Software AS, Norway). Two different ways of relating PCA to consumer attributes will be handled here.

OPTION 1: Relating PCA components to the consumer attributes

In this case the principal components (PCs) are independently related to consumer attributes (here external variables) using simple PLS regression (see Section 2.2 for arguments).

OPTION 2: Individual preferences and consumer segmentation

In this case, a visual segmentation based on the first PCA score is performed and used for illustration of the method. Visual segmentation is sometimes more relevant than using a clustering algorithm since there are usually no clear segments in this type of studies (Næs, et al., 2010, Endrizzi et al., 2011). In a visual approach, segmentation can be done according to the interpretation that one is interested in investigating in more detail. Finally, consumers are characterized in terms of socio-demographics, attitudes and habits with the help of a PLS-DA regression model relating the defined segments to the questionnaire.

Note that since this approach is based on the same basic data as for Option 1, one can in many cases not expect large differences in conclusions between the two options. Option 2 is, however, more specific in the sense that it can also be used for segments with a special shape not directly related to one of the components which is the case for the one used below for illustration purposes.

We refer to Section 2.2 for a more detatiled analysis of how the PLS regression method was used.

4. RESULTS

4.1 STRATEGY 1: Simultaneous Mixed Logit Model of the conjoint factors and consumer attributes (One-step strategy)

Table 3 contains the estimated parameters of the Mixed Logit model (means and standard deviations) for the main effects of the conjoint factors, their interactions and interactions with sociodemographics terms at population level as well as the variability of the individual coefficients as measured by SD. The null hypothesis that all coefficients are zero is rejected by a Wald test (p-value <0.001) which indicates that the attributes chosen are considered relevant by consumers. The number of observations in the model is equal to 2376, which corresponds to n = 99 participants and not n = 102, because three consumers did not declare their age.

Note that the results are slightly different from the results in paper (Asioli et al., 2016) for the same data. The reason for this is that the methods is iterative and that in the present article we used 4,000 so-called halton draws instead of 2,000 in the previous paper (Asioli et al., 2016). As can be seen, however, the p-values for the different tests are quite similar to each other and none of the general conclusions is altered.

Table 3 – Estimated parameters for ML model with conjoint variables' main effects and interactions, and interactions with socio-demographic attributes (Strategy 1). The two columns to the left refer to the population effects while the two columns to the right correspond to the individual differences as measured by standard deviations (SD).

<<Please, place here table 3>>

On average the consumers prefer low calorie coffees, Norwegian origin and low prices while they do not seem to have any strong differences in preference for the two Coffee types (Table 3). However, Price has a stronger negative effect than Origin and Calories. It is interesting to note that only main effect Coffee type has significant SDs (see Asioli et al., 2016 for more details), indicating large individual differences in preference for this factor. In other words, even without a significant overall effect of coffee, there is a lot of individual variation among consumers.

With regard to the interaction effects among conjoint factors the only significant interaction effect (in the population) detected is Coffee*Price (p=0.012) (Table 3). Thus, consumers who prefer latte are a little bit more sensitive to price changes than consumers who prefer espresso, showing a slightly stronger preference for low price. With regard to the interaction effects crossing conjoint factors with socio-demographic attributes, the most significant interaction effects are Calories*Gender (p<0.001) and Coffee*Gender (p=0.034) (Table 3). This indicates that males and females (on average) show different preferences for calorie contents and iced coffee types (i.e. Latte and Espresso). More specifically, females prefer low calories much more strongly than males. Interaction plots illustrating these results are available in Asioli et al. (2016)[1].

It is interesting to note that there are several interaction effects (i.e. Coffee*Calories, Coffee*Age, Origin*Age, Price*Age) with significant standard deviations (SDs), indicating the relevance of individual differences and also differences within the genders and age groups that are not visible when looking only at the average Gender and Age effects.

4.2 STRATEGY 2: Mixed Logit Model, PCA, PLS regression and PLS discrimination (Multi-step strategy)

4.2.1 Mixed Logit Model

Table 4 contains the estimated parameters of the Mixed Logit model (means and standard deviations) for the main effects of the conjoint factors and their interactions terms at

[1] As indicated before, the model used here is a bit different (different number of iterations), but the results are similar as well as the interaction plots.

population level as well as as the variability of the individual coefficients as measured by SD. Again the null hypothesis that all coefficients are zero is rejected by a Wald test (p-value <0.01).

Table 4 – Estimated parameters for ML model with conjoint variables' main effects and interactions (Strategy 2). The two columns to the left refer to the population effects while the two columns to the right correspond to the individual differences as measured by standard deviations (SD).

<<Please, place here table 4>>

From Table 4 we can see again that on average consumers prefer low calories, low prices and Norwegian origin while coffee type is not significant at mean population level which is consistent with results obtained from strategy one (see section 4.1.1). It is interesting to note that all the conjoint factors (main effects) have significant standard deviations (SDs) meaning that there are individual differences in perception. This corresponds to the results in strategy one with significant SD's for several of the interactions with Gender and Age. But as can be seen, in this case without Age and Gender effects, this element appears in the SD's for the main effects themselves. In strategy two these individual differences will be further investigated in the following steps.

From Table 4 we can see that only one interaction is significant, namely the interaction between coffee type and price (Coffee*Price), again corresponding to above.

4.2.2 Principal Component Analysis (PCA) on regression coefficients

In order to further investigate consumer attributes, a PCA model was run on individual regression coefficient estimates from the ML model above (i.e. model including only main effects and interactions of conjoint factors) (Figure 2). In the PCA model the coefficients are not standardized to preserve the original scale variations. In the following, we concentrate on four principal components (PCs), corresponding very well with the four design factors in the following order: Coffee type (on PC-1, explaining 86% of the variance), Origin (on PC-2, explaining 6% of the variance), Calories (on PC-3, explaining 4% of the variance) and Price (on PC-4, explaining 3% of the variance). The correspondence between principal components and design factors is natural because of the orthogonality of the design. As can also be seen, the order of importance does not match the relative importance of the factors at a population level (averages) indicated in the ML model, while it corresponds very well with the order indicated by the significant SD's in Table 4. Thus, it is clear that Coffee type explains the largest variance, followed by Origin and Calories. It is also interesting to note that Price contributes least to the variance. This is because there is a strong agreement between consumers in the direction of preferring a lower price for the same product attributes. On the contrary, there is no preferred type of coffee at population level (this main effect is non significant), but a lot of individual variations revealed by the SDs and the PCA results. This clearly shows the shortcomings of only looking at average effects that is often done in many conjoint studies.

It is important to emphasize that instead of the PCs of the regression coefficients one could in this case, based on an orthogonal design, have used the main effect estimates for the consumers directly as response variables. For non-orthogonal designs, the relation between main effects and the PCA plot may be more complicated. Using the PCA also opens up the possibility of identifying more easily consumers with for instance large values on two or more of the components. This latter aspect could be important for segmentation purposes as is the case for the Option 2 below.

Figure 2 – PCA correlation loadings plot - for PC-1 and PC-2 - on individual Mixed Logit parameter estimates from choice data (scores are presented in Figure 6)

<<Please, place here figure 2>>

Note: the names placed in the figure on the extremes of PC-1 (Espresso and Latte) and PC-2 (Italy and Norway) have been inserted for a better interpretation of the bi-plot.

4.2.3 Investigation of consumer attributes

As indicated in the section 3.5.2 two options for investigating consumer attributes starting from the PCA analysis will be tested. The first option relates consumer attributes as external variables directly to the PCs indentified using for instance PLS regression, while in the second option the consumer attributes are related to segments determined in the PCA plot, using PLS-DA. In all cases, the PLS regression allows for many collinear explanatory attributes which is a clear advantage of the method. The values of the explained variances indicated in the next steps refer to the plots with only significant consumer attributes.

OPTION 1: Relating PCs to consumer attributes

We applied PLS regression by relating the PCs identified in the PCA above directly to consumer attributes. Due to the independence of the axes, it is most natural here to consider the axes separately (individual PCs), but a joint analysis is also possible (see above). The results from components 3 and 4 will only be mentioned briefly without Figures.

Figure 3 presents PC-1 (Coffee type) and its relation to consumer attributes. The cross-validation (CV) indicates that one component is clearly significant, but component two also added slightly to prediction ability. The explained variances for components 1 and 2 are equal to 20% and 11% for X and 50% and 5% for Y. We can notice that there is a large number of significant, as determined by the jack-knife method described above for 1 component,consumer attributes as compared to the other PCs (see for instance Figures 4 for PC-2 results).

In particular, PC-1 (describing conjoint factor Coffe type, see Figure 3) is positively correlated to espresso coffee habits (preference for high coffee intensity, warm coffee, espresso, americano, regular and black coffee) and males while it is negatively correlated to consumption habits of warm coffee with milk (e.g. milk content, latte and cappuccino) (Table 5). Thus PC-1 describes two directions of coffee type habits, which also indicates the possibility to identify two groups of consumers as we will see in the option two. As can be seen, there is a natural correspondence between the preference pattern and what the consumers indicate that they do/like. The position of the consumer attributes in the plots before and after the significant test is more or less the same in both configurations.

Gender

Figure 3 – Correlation loadings - PLS components 1 and 2 - with significant consumer attributes from PLS regression model using PC-1 as dependent variable (Coffee type)

<<Please, place here figure 3>>

Using two components in the significance tests changed the number of significant attributes slightly. In particular, two attributes related to iced coffee habits (preference for brand B and canteen as location of iced coffee consumption) have now a significantly positive correlation to PC-1 (Coffe type direction). On the other hand preference for Brand A iced coffee, americano warm coffee and indication of work/university as usual location of warm coffee consumption are no longer significant. All attributes that are significant for both one and two components PLS models are located in the same positions in both plots. For two components Gender was not significant, but this is not so surprising since Gender is only borderline significant in Strategy 1.

Table 5 – Significant consumers attributes for the one-component model (PC1) (p-values on regression coefficients, from jack-knife test)

<<Please, place here table 5>>

For PC-2, the predictive CV indicated that none of the components was significant, but based on one component the jack-knife significance test gave a number of significant attributes. Figure 4 shows the relation of PC-2 (describing conjoint factor Origin, see Figure 4) with significant consumer attributes. The explained variances for 1 and 2 components are now 36% and 16% for X and 21% and 1% for Y.

Figure 4 – Correlation loadings - PLS components 1 and 2 - with significant consumer attributes from PLS regression model using PC-2 (Origin)

<<Please, place here figure 4>>

We can see that PC-2 is positively related to location of iced coffee consumption (i.e. café/restaurant and bar) which is negatively correlated to consumer attributes importance of origin and preference for foods of Norwegian origin and for familiar foods (Table 6). Neither Age nor Gender were significant in this case, which corresponds to the findings from Strategy 1. The position of the consumer attributes in the plots before and after the significant test and variable selection is more or less the same.

Table 6 – Significant consumers' attributes for the two-component model (PC1-PC2) (p-values on regression coefficients, from jack-

<<Please, place here table 6>>

For PC-3 (describing conjoint factor Calories) the cross-validation (CV) indicates a slight significance of the first component and therefore only one component was used in the jack-knife test. PC-3 was found to be positively correlated with price and Gender (males) and negatively correlated with calories, use of sweetener and warm coffee habits (i.e. cappuccino and americano). Gender was in this case one of the significant attributes which is positively correlated to PC-3. This indicates that the differences between the calorie levels is more important for the females than it is for the males (Asioli et al., 2016), which is in correspondance with the results for Strategy 1. The position of the consumer attributes in the plots before and after the significant test is more less the same. Finally, PC-4 (which is related to individual differences in perception of price) is positively correlated to origin and negatively correlated to price and calories. Again no component was significant in the cross-validation, and only one component was used in the jack-knife test. The attributes reported here are the ones found to be significant. In this case neither Age nor Gender was significant. The position of the consumer attributes in the plots before and after the significance test and variable selection is more or less the same.

As we have seen, in these analyses, Gender shows up as significant for PC-1 and PC-3 (i.e. for coffee and calories). This means that the two genders have a different preference for the two coffee types and calories levels, i.e. there is an interaction between the two. This corresponds exactly to what was found in Strategy 1 where the interaction between Gender and the two conjoint factors (coffee type and calories) were the only two interactions found to be significant (see Table 3). In the present Strategy (option one), however, one can also obtain information about the other attributes and how they relate to the conjoint factors which is clearly more difficult in Strategy 1.

Quantification of the individual differences in the interactions between Gender and conjoint factors which was a major issue in the previous strategy is, however, less obvious in the present case. One can see clear individual differences in the scores plot regarding preferences along the different conjoint factors, but a numerical statement of significance is not available here, in contrast to Strategy 1.

Note that for none of the analyses the significance tests and elimination of the non-signifiant variables changed the general structure/position of the reminaing variables. The elimination of variables must here therefore mainly be considered a way of making plots interpretation simpler.

OPTION 2: Preference heterogeneity and consumer segmentation

Espresso and Latte segments (PCA)

For comparison with the above and for illustrating this second option we decided to concentrate on two equally-sized segments determined along the first PCA axis. It should, however, be emphasised that other PCs can be used to define segments depending on the objective of the study. For example, four segments defined along PC1 and PC2 could also be used as has been done in a previous paper with rating data (Asioli et al., 2014). Indeed, visual segmentation can easily be performed and it is flexible (Almli et al., 2015; Næs, et al., 2010). The consumer segments consist of 51 subjects for the Espresso group and 51 subjects for the Latte group (Figure 5). In the following sections these segments are referred to as "Espresso" and "Latte" segments, respectively (see section 4.2.2).

Figure 5 – PCA scores plot on individual Mixed Logit parameter estimates from choice data

<<Please, place here figure 5>>

Segments characteristics

To describe the consumer segments in terms of habits, attitudes and socio-demographic attributes an approach based on PLS-DA was used (Figure 6). The consumer groups (Latte and Espresso) were represented by dummy variables (Ys) in the PLS-DA, while consumer attributes were used as independent variables (Xs). The cross-validation (CV) indicates that only one component had a significant prediction ability and therefore only one component was used in the jack-knife test. The explained variances for the first two components were 29% and 19% for X and 34% and 1% for Y. Socio-demographic attributes were not found to be significant.With regard to warm coffee consumption habits, the two segments differ significantly for several attributes. Consumers in the Espresso group show the highest consumption of “Espresso” warm coffee type and also preference for “black” warm coffee. Finally, consumers belonging to the Latte segment have preference for two types of warm coffee: latte and capuccino. These findings are fully coherent with the definition of the two groups. Further, only one iced coffee habit has been found significant which is the preference for Espresso segment of the “B” brand.

As can be seen these results are similar to the results of the PC-1 in the option one which is natural since we segmented the consumers based on PC-1. The main reason for incorporating the Option 2 here, however, is that it can also be used for other segments with shapes and positions not directly related to one of the components as was the case here.

As can be seen Gender is no longer significant at the fixed significance level. As discussed above this is not totally surprising since Gender is borderline significant and therefore two different tests may lead to different conclusions relative to a fixed significance threshold.

Figure 6 – Correlation loadings with significant consumer attributes from PLS-DA model

<<Please, place here figure 6>>

5. DISCUSSION

The main aim of this paper was to compare two different strategies for investigating individual differences among consumers using choice data collected in a study about consumer preferences for iced coffee products in Norway. The focus of the paper is on methodology and advantages and disadvantages from a methodological point of view. It must, however, be emphasized that the methods should be compared on more data sets in order to come with more general statements about their properties.

5.1 Comparison of the two strategies in terms of flexibility

The *multi-step* Strategy (here Strategy 2) can be considered more flexible compared to the *one-step* Strategy (here Strategy 1) since the latter is only able to investigate a limitated number of pre-defined consumer attributes at a time. The *multi-step* Strategy on the other hand can be used to investigate a large number of potentially collinear consumer attributes at the same time. This is important since no selection of attributes is needed before analysis. Options 1 and 2 for Strategy 2 are more or less equally flexible. For the first one, one can relate the regression coefficients or their PCs as done here directly to the consumer attributes, while for option 2 one can look at different segments depending on the scope of the analysis. The latter then opens up for a more focused analysis related to what one is most interetested in studying.

5.2 Comparison of the two strategies in terms of data analysis, computation and interpretation

Data analysis and computation of the *one-step* strategy can be considered simpler to perform compared to the *multi-step* strategy. First of all the *one-step* strategy requires skills and expertise related to only one statistical model (Mixed Logit Model) while in the *multi-step* strategy three models have to be performed (Mixed Logit Model, PCA and PLS regression). This also means that it may require expertise and skills about two software programs, such as (in this case) STATA 11.2 and The Unscrambler X 10.2.

For the comparison of options 1 and 2 for Strategy 2, the second one is more complex since an additional step of choosing the segments comes in on top of ML modelling and regression.

From an interpretation point of view, Strategy 1 is slightly simpler since all results are to be found in one table only. However Strategy 2 has the advantage of using maps which are very easy to understand in comparison with estimate values, especially for non statisticians.

5.3 Comparison of the two approaches in terms of outcomes

A possible drawback with Strategy 2 is that it is harder to obtain quantitative information about the individual differences in consumers' liking for a conjoint factor within for instance a consumer attribute such as Gender or Age. It may be visible in the plot that such a tendency is clear, but a quantitiave assessment is more difficult to get.

For the elements that can be compared the two strategies led in this case to similar results regarding the main and interaction effects among the conjoint factors. Indeed, both strategies show that consumers have strong preferences for low calories, Norwegian origin and low price iced coffee products as main effects, while there is a significant effect for the interaction Coffee*Price. Strategy 2, however, added information about a number of other consumer attributes which may be very important for product development practices.

6. CONCLUSIONS

This study compared two different ways investigating individual differences and their relation to consumer attributes using choice data. One of the strategies is a *one-step* a priori segmentation strategy based on joint Mixed Logit modelling of all data. The other strategy is a *multi-step* strategy based on relating the individual preference results from the Mixed Logit model to the external consumer attributes by regression or classification methods. Outcomes showed that the two strategies for the actual data gave similar results about main and interaction effects among conjoint factors. For the individual differences, the results were also comparable for the consumer attributes that were considerd in both strategies. The *multi-step* strategy has the advantage that it is more flexible and can be used to analyse several, possibly collinear, consumer attributes at the same time. An advantage of the *one-step* strategy is that it gives simpler numerical assessments of individual differences in their assessments of the different conjoint factors. On the other hand, it only allows to focus on few pre-selected consumer attributes. 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. Comparisons of the two methodologies for other data sets are needed in order to evaluate the general validity of the conclusions.

7. ACKNOWLEDGEMENTS

The research was conducted in Norway in the framework of the FOODCHOICE project, a user-driven project funded by the Research Council of Norway focusing on method development within consumer research. Financial support from the Research Council of Norway for the YGGDRASIL mobility program 2012–2013: Project Number 219787/F11 is acknowledged. Thanks also to the European Commission through the Marie Curie Actions Intra European Fellowship (IEF), call FP/-PEOPLE-I2012-IEF – project title ‘‘Innovative Methodologies for New Food Product Development: combining Sensory Science and Experimental Economics – NEFOMET” for the support. Kjell Merok (Nofima) is thanked for designing the iced coffee products profiles.

8. REFERENCES

Alfnes, F. (2004). Stated preferences for imported and hormone-treated beef: application of a mixed logit model. *European Review of Agricultural Economics* , *31* (1), 19–37.

Almli, V. L., Øvrum, A., Hersleth, M., Almøy, T., & Næs, T. (2015). Investigating individual preferences in rating and ranking conjoint experiments. A case study on semi-hard cheese. *Food Quality and Preference*, *39*, 28–39.

Asioli, Næs, T., Øvrum, A., & Almli, V. L. (2016). Comparison of rating-based and choice-based conjoint analysis models. A case study based on preferences for iced coffee in Norway. *Food Quality and Preference*, *48*, 174–184.

Asioli, D., Næs, T., Granli, B. S., & Lengard Almli, V. (2014). Consumer preferences for iced coffee determined by conjoint analysis: an exploratory study with Norwegian consumers. *International Journal of Food Science & Technology*, *49*(6), 1565–1571.

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, *17*(3), 166–173.

Bech, M., & Gyrd-Hansen, D. (2005). Effects coding in discrete choice experiments. *Health Economics*, *14*(10), 1079–1083.

Bonnet, C., & Simioni, M. (2001). Assessing consumer response to Protected Designation of Origin labelling: a mixed multinomial logit approach. *European Review of Agricultural Economics* , *28* (4), 433–449.

Boxall, P. C., & Adamowicz, V. (2002). Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach. *Environmental and Resource Economics*, *23*(4), 421–446.

Carson, R., Louviere, J., Anderson, D., Arabie, P., Bunch, D., Hensher, D., … Wiley, J.

(1994). Experimental analysis of choice. *Marketing Letters*, *5*(4), 351–367.

Combris, P., Bazoche, P., Giraud-Héraud, E., & Issanchou, S. (2009). Food choices: What do we learn from combining sensory and economic experiments? *Food Quality and Preference*, *20*(8), 550–557.

Endrizzi, I., Gasperi, F., Rødbotten, M., & Næs, T. (2014). Interpretation, validation and segmentation of preference mapping models. *Food Quality and Preference*, *32*, 198–209.

Endrizzi, I., Menichelli, E., Johansen, S. B., Olsen, N. V., & Næs, T. (2011). Handling of individual differences in rating-based conjoint analysis. *Food Quality and Preference, 22*(3), 241–254.

Enneking, U., Neumann, C., & Henneberg, S. (2007). How important intrinsic and extrinsic product attributes affect purchase decision. *Food Quality and Preference*, *18*(1), 133–138.

Green, P. ., & Srinivasan, V. (1990). Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice. *Journal of Marketing*, *54*(4), 3–19.

Gustafsson, A., Herrmann, A., & Huber, F. (2003). *Conjoint measurement: Methods and applications*. Berlin: Springer.

Haaijer, R., Kamakura, W. A., & Wedel, M. (2001). The “No-Choice” Alternative in Conjoint Choice Experiments. *International Journal of Market Research*, *43*(1), 93–106.

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal*, *7*(3), 188–401.

Johanna Lena, H., & Roosen, J. (2015). Consumer Heterogeneity in the Willingness to Pay for Local and Organic Food. *Journal of Food Products Marketing*, *21*(6), 608–625.

Louviere, J. J., Hensher, D., & Swait, J. (2000). *Stated choice methods: analysis and* *application*. Cambridge.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods. Analysis and applications*. Cambridge University Press.

Lusk, J. L., Fields, D., & Prevatt, W. (2008). An Incentive Compatible Conjoint Ranking Mechanism . *American Journal of Agricultural Economics* , *90* (2), 487–498.

Martens, H., & Martens, M. (2000). Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Quality and Preference*, *11*(1–2), 5–16.

Molteni, L., & Troilo, G. (2007). *Ricerche di Marketing*. Milano: McGraw-Hill.

Næs, Brockhoff, P., & Tomic, O. (2010). Statistics for sensory and consumer science.

Næs, T., Almli, V. L., Bølling Johansen, S., & Hersleth, M. (2010). Alternative methods for combining design variables and consumer preference with information about attitudes and demographics in conjoint analysis. *Food Quality and Preference*, *21*(4), 368–378.

Næs, T., Kubberød, E., & Sivertsen, H. (2001). Identifying and interpreting market segments using conjoint analysis. *Food Quality and Preference*, *12*(2), 133–143.

Ståhle, L., & Wold, S. (1987). Partial least squares analysis with cross-validation for the two-class problem: A Monte Carlo study. *Journal of Chemometrics*, *1*(3), 185–196.

Train, K. (2009). *Discrete choice methods with simulation*. (C. U. Press, Ed.)*New York* (Vol. 47). Cambridge University Press.

Van Loo, E. J., Caputo, V., Nayga, R. M., Meullenet, J.-F., & Ricke, S. C. (2011). Consumers' willingness to pay for organic chicken breast: Evidence from choice experiment. *Food Quality and Preference*, *22*(7), 603–613.

Van Wezemael, L., Caputo, V., Nayga, R. M., Chryssochoidis, G., & Verbeke, W. (2014).

European consumer preferences for beef with nutrition and health claims: A multi-country investigation using discrete choice experiments. *Food Policy*, *44*, 167–176.

Vigneau, E., Endrizzi, I., & Qannari, E. M. (2011). Finding and explaining clusters of consumers using the CLV approach. *Food Quality and Preference*, *22*(8), 705–713.

Wold, S., Martens, H., & Wold, H. (1983). Matrix Pencils: Proceedings of a Conference Held at Pite Havsbad, Sweden, March 22--24, 1982. In B. Kågström & A. Ruhe (Eds.), (pp. 286–293). Berlin, Heidelberg: Springer Berlin Heidelberg.

Yangui, A., Akaichi, F., Costa-Font, M., & Gil, J. M. (2014). Are ranking preferences information methods comparable with the choice experiment information in predicting actual behaviour? In *EAAE 2014 Congress "Agri_Food and Rural Innovations for Healthier Societies"* (p. 13). Ljubljana, Slovenia.

Øvrum, A., Alfnes, F., Almli, V. L., & Rickertsen, K. (2012). Health information and diet choices: Results from a cheese experiment. *Food Policy*, *37*(5), 520–529.

Appendix I – Choice design

"Please, place here appendix I"

Highlights

- Two strategies investigating individual differences using choice data are compared.
- Strategy 1 includes the consumer attributes directly in the Mixed Logit Model.
- Strategy 2 combines different methods such as Mixed Logit Model, PCA and PLS.
- Strategy 2 is preferred for its flexibility and stronger exploratory capabilities.

Table 1 – Conjont factors and levels used in the conjoint design

FACTORS	LEVELS	
Coffee type	1	Latte
	2	Espresso
Calories	1	60 kcal/100 ml
	2	90 kcal/100 ml
Origin	1	Norway
	2	Italy
Price	1	17 NOK (≈ € 2.0)
	2	23 NOK (≈ € 2.7)
	3	29 NOK (≈ € 3.4)

Table 2 – Factors coded and their description

FACTOR	DESCRIPTION
Coffee	If Espresso: 1; otherwise (Latte): -1
Calories	If 90 kcal/100 ml: 1; otherwise (60 kcal/100 ml): -1
Origin	If Italy: 1; otherwise (Norway): -1
Price	If 17 NOK: -1; if 23 NOK: 0; if 29 NOK: 1
Gender	If Male: 1; otherwise (Female): -1
Age	If age is 37-56: 1; otherwise 21-36 (younger): -1

Table 3 – Estimated parameters for ML model with conjoint variables' main effects and interactions, and interactions with socio-demographic attributes. The two columns to the left refer to the population effects while the two columns to the right correspond to the individual differences as measured by standard deviations (SD).

EFFECTS	**GROUP AVERAGE**		**INDIVIDUAL VARIATION**	
	Estimate	**p-Value**	**Std. Dev**	**p-Value**
Main effects				
Coffee	-0.046	0.883	2.463	0.000***
Calories	-0.657	0.000***	0.317	0.232
Origin	-0.500	0.005**	0.152	0.468
Price	-1.696	0.000***	0.181	0.462
Interactions among conjoint factors				
Coffee*Calories	-0.046	0.737	0.526	0.005**
Coffee*Origin	0.298	0.093	0.477	0.051
Coffee*Price	0.316	0.012*	0.008	0.947
Calories*Origin	0.085	0.526	0.007	0.962
Calories*Price	-0.016	0.907	0.268	0.274
Origin*Price	-0.113	0.454	0.276	0.237
Interactions with sociodemographics attributes				
Coffee*Gender	0.569	0.034*	0.918	0.063
Coffee*Age	-0.492	0.057	1.310	0.001**

Calories*Gender	0.544	0.000***	0.105	0.648
Calories*Age	-0.144	0.258	0.660	0.001*
Origin*Gender	0.075	0.661	0.281	0.170
Origin*Age	0.144	0.391	1.136	0.000***
Price*Gender	-0.127	0.467	0.510	0.047*
Price*Age	0.271	0.130	0.991	0.000**

*, ** and *** indicate significant effects at 0.05, 0.01 and 0.001 levels, respectively.

Number of choice observations: 2376

Number of consumers: 99

Table 4 – Estimated parameters for ML model with conjoint variables' main effects and interactions. The two columns to the left refer to the population effects while the two columns to the right correspond to the individual differences as measured by standard deviations (SD).

EFFECTS	**GROUP AVERAGE**		**INDIVIDUAL VARIATION**	
	Estimate	**p-Value**	**Std. Dev**	**p-Value**
Main effects				
Coffee	-0.183	0.379	1.881	0.000***
Calories	-0.571	0.000***	0.557	0.000***
Origin	-0.281	0.007**	0.666	0.000***
Price	-1.06	0.000***	0.596	0.000***
Interactions among conjoint attributes				
Coffee*Calories	0.061	0.537	0.204	0.393
Coffee*Origin	0.162	0.203	0.306	0.235
Coffee*Price	0.229	0.015*	0.007	0.949
Calories*Origin	0.046	0.676	0.042	0.711
Calories*Price	-0.062	0.500	0.073	0.752
Origin*Price	-0.111	0.335	0.052	0.763

*, ** and *** indicate significant effects at 0.05, 0.01 and 0.001 levels, respectively.

Number of choice observations: 2448

Number of consumers: 102

Table 5 – Significant consumers attributes for the one-component model (PC1) (p-values on regression coefficients, from jack-knife test)

CONSUMERS ATTRIBUTES	P-VALUES
Coffee intensity	0.000
Warm Coffee	0.001
Tine IC	0.038
Regular C	0.000
Latte C	0.000
Espresso C	0.000
Capp. C	0.020
Mocca C	0.015
Americano C	0.017
Black	0.000
Milk	0.001
Work/Un C	0.019
Gender	0.040

Table 6 – Significant consumers' attributes for the two-component model (PC1-PC2) (p-values on regression coefficients, from jack-knife test)

CONSUMERS ATTRIBUTES	P-VALUES
Origin	0.027
Late at night	0.049
Café'/restaurant	0.029
Bar IC	0.026
Best food own	0.000
Stick foods	0.002
Norwegians	0.000

Appendix I – Choice design

SET	COFFEE TYPE	CALORIES (kcal per 100 ml)	ORIGIN	PRICE (NOK)
1	Espresso	90	Italy	23
	Latte	60	Norway	17
	Latte	90	Norway	29
2	Latte	90	Italy	29
	Latte	90	Italy	17
	Espresso	60	Norway	23
3	Espresso	60	Norway	29
	Latte	60	Italy	17
	Latte	90	Norway	23
4	Espresso	90	Norway	29
	Espresso	60	Italy	23
	Latte	60	Italy	17
5	Espresso	60	Norway	17
	Latte	60	Italy	29
	Latte	90	Italy	23
6	Latte	60	Norway	29
	Espresso	90	Norway	17
	Espresso	60	Italy	23
7	Latte	90	Norway	23

<table>
<tr><td rowspan="2"></td><td>Espresso</td><td>90</td><td>Italy</td><td>17</td></tr>
<tr><td>Espresso</td><td>60</td><td>Italy</td><td>29</td></tr>
<tr><td rowspan="3">8</td><td>Latte</td><td>60</td><td>Norway</td><td>23</td></tr>
<tr><td>Espresso</td><td>90</td><td>Italy</td><td>29</td></tr>
<tr><td>Espresso</td><td>90</td><td>Norway</td><td>17</td></tr>
</table>

Figure 1 – One of the iced coffee profiles

Figure 1 – One of the iced coffee profiles

919

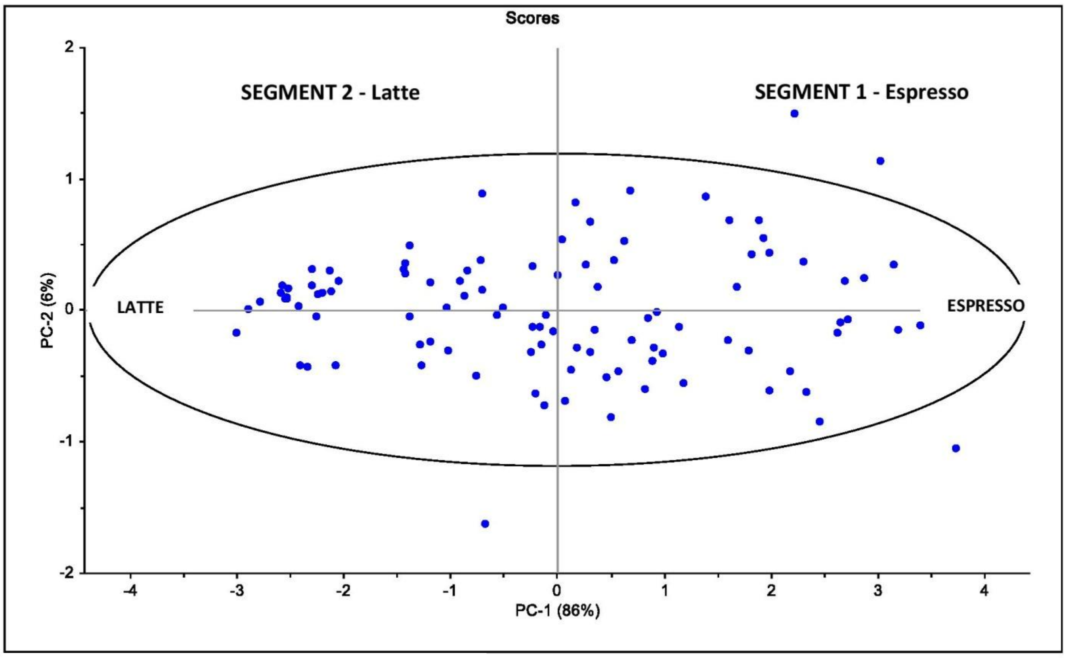
Scores
SEGMENT 2 - Latte
SEGMENT 1 - Espresso
LATTE
ESPRESSO
PC-2 (6%)
PC-1 (86%)
2
1
0
-1
-2
-4
-3
-2
-1
0
1
2
3
4

Figure 2

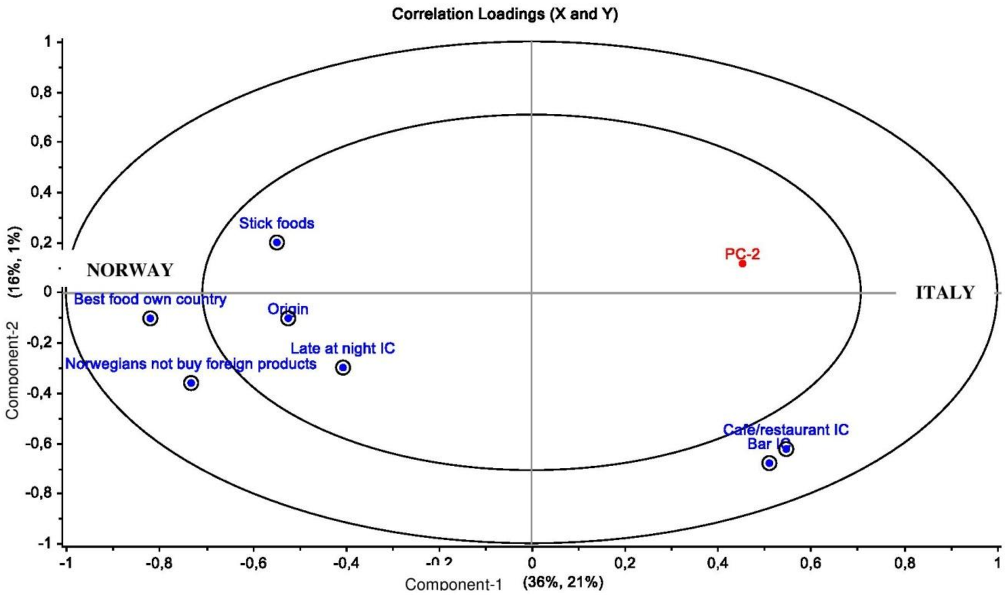
Correlation Loadings (X and Y)
Stick foods
PC-2
NORWAY
ITALY
Best food own country
Origin
Late at night IC
Norwegians not buy foreign products
Cafe/restaurant IC
Bar IC
Component-2 (16%, 1%)
Component-1 (36%, 21%)

Figure 3

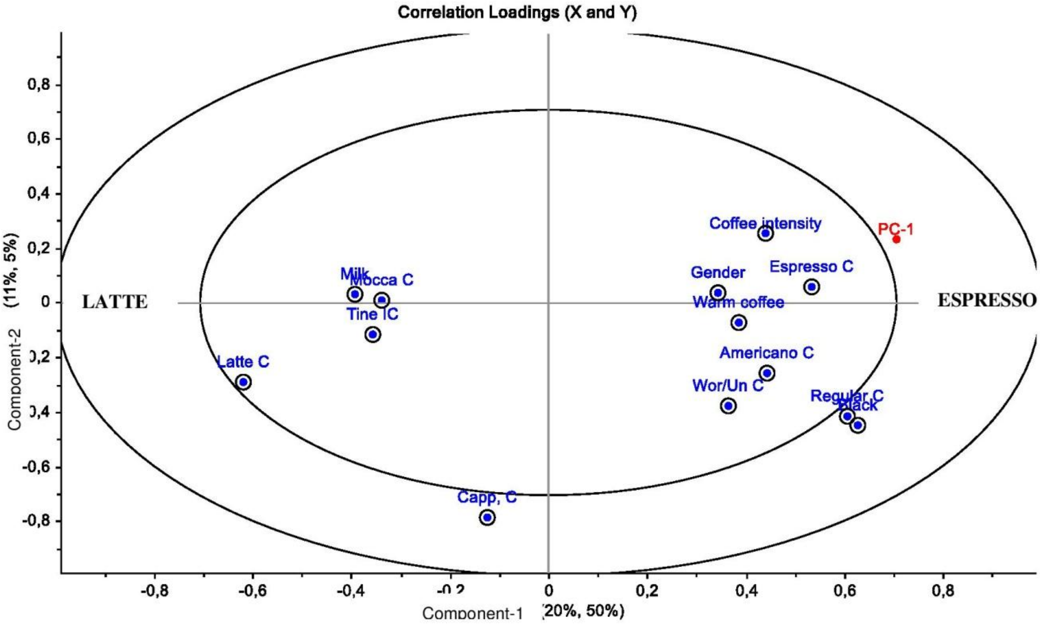
Correlation Loadings (X and Y)
LATTE
ESPRESSO
Coffee intensity
PC-1
Gender
Espresso C
Warm coffee
Milk
Mocca C
Tine IC
Latte C
Americano C
Wor/Un C
Regular C
Black
Capp, C
Component-2 (11%, 5%)
Component-1 (20%, 50%)

Figure 4

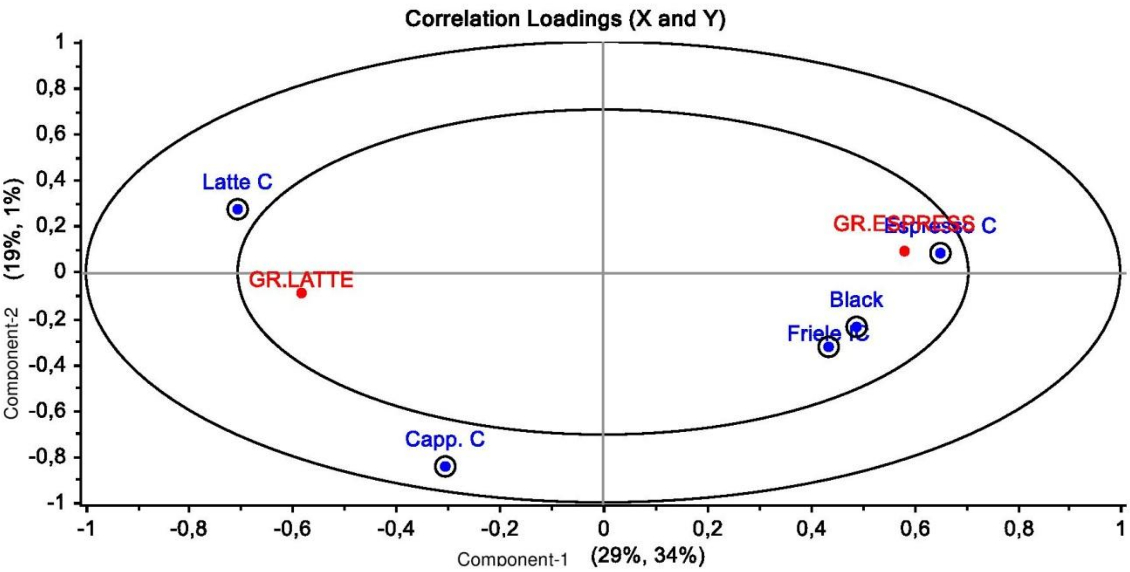
Correlation Loadings (X and Y)
Latte C
GR.LATTE
GR.ESPRESS
Espresso C
Black
Friele IC
Capp. C
1
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
-0,8
-1
-1
-0,8
-0,6
-0,4
-0,2
0
0,2
0,4
0,6
0,8
1
Component-2 (19%, 1%)
Component-1 (29%, 34%)

Figure 5