MaxDiff approaches for PDO "Calanda" peaches (Spain)

Etiénne Groot[1], Luis M. Albisu

[1] Centre of Agrofood Research and Technology of Aragón (C.I.T.A.), Zaragoza, Spain
(egroot@aragon.es, lmalbisu@aragon.es)

Paper prepared for presentation at the 113th EAAE Seminar "A resilient European food industry and food chain in a challenging world", Chania, Crete, Greece, date as in: September 3 - 6, 2009

MaxDiff approaches for PDO “Calanda” peaches (Spain)

Etiénne Groot[1], Luis M. Albisu

[1] Centre of Agrofood Research and Technology of Aragón (C.I.T.A.), Zaragoza, Spain
(egroot@aragon.es, lmalbisu@aragon.es)

Abstract. Peaches with PDO Calanda are one of the 20 fruits with PDO existing in Spain. The aim of this work is to understand how consumers make their choices based on the most important peaches’ attributes and levels. In this work, 4 attributes with 3 levels in each attribute have been considered (price: 1.5 €/kg, 2.5 €/kg and 3.5 €/kg; origin: PDO Calanda, non PDO Calanda and non Calanda; packaging: bulk, conventional packaging and active packaging; and fruit size: small, medium and big). Four Best-Worst (BW) exploded models have been utilised, two of them with scale factors. All those models have been compared to two traditional Discrete Choice (DC) models. Results show that traditional DC models have better performance than the other models and the best model is when consumers select the best option. Within the exploded models, the choice sequence decisions starting from worst options are better than those which start from best option. Consumers prefer PDO Calanda peaches over other types and the positive difference in their Willingnes to Pay (WTP) are more or less the same between peaches from Calanda with PDO and without PDO as it is between the latter and peaches coming from other origins.

Keywords: best-worst, exploding models, consumer behavior, choice experiment, fruit quality.

1. Introduction

Origin labels have been promoted in the European Union to increase and to promote the quality of food products coming from specific geographic locations. In Spain, there are 288 Protected Designation of Origin (PDO) and Protected Geographical Indication (PGI) food products and 20 of them are fruits. Peaches with PDO Calanda are the only ones with EU quality label recognition in Spain. Calanda is located in the east part of Aragon and its climatology is adequate to produce high quality peaches, which are sold at the end of the production season, covering from the beginning of September until the end of October. Peaches from Calanda have had high market recognition for its quality during many years. In 1999 the Regulatory Council of the Protected Designation of Origin Calanda peaches was established with the objective of assuring the product quality and maintaining its reputation.

Growers and commercial companies must comply the rules established by the Regulatory Council to get the PDO label (BOA, 1999). Jesca, Evaisa and Calante are the only varieties accepted as PDO Calanda peaches and they have to be produced in the Calanda area, which includes 44 municipalities. Growers employ techniques as the “aclareo” and the fruit “bag production”. The “aclareo” consists of taking off 70% of the fruits at their first development stage and, as a consequence, fruits get much bigger. The “bag production” fruit technique is necessary to avoid the Mediterranean fly attack. Both operations demand half of the labour force, which represents one quarter of the total production cost (Mainar, 2006).

At harvest, the minimum fruit size has to be about 73 mm, which corresponds to calibre AA, with flash hardness between 3.5 and 5.0 kg/ 0.5 cm^2 and sugar content superior to 12º Brix. Mature fruits are yellow coloured, without green or orange tonalities, showing that they have reached a good ripening level. Hurt peaches or those with any injury are forbidden to be certified.

A minimum quality standard on appearance and taste is important to guarantee peach purchasing and consumer loyalty (Predieri *et al*., 2006; Crisosto *et al*., 2005; Bruhn, 1994). The presence of brands allows consumers’ product identification and more accurate quality. Consumers’ quality perception increases brand value. Polo (2007) found that the PDO Calanda peaches brand was valued by Madrid and Barcelona wholesalers. Results showed that 40% of them recognised that peaches with PDO have prices 20% greater than the same peaches without PDO. This work deals with how consumers evaluate PDO Calanda peaches.

The Discrete Choice Experiment (DCE) has been employed with the best-worst scaling approach. In best-worst experiments respondents need to choose best and worst options in a set of alternatives. An assumption of this process is that respondents have the same ability to state best and worst options. A subsidiary objective of this work is to check whether there are differences when respondents decide best and worst options. The next section describes the different Discrete Choice models, with a special emphasis in best-worst scaling, providing analytical theoretical support. The following section refers to the experimental design. Section 3 presents the results and the final remarks can be found in the last section.

2. Methodological approaches

2.1. Discrete Choice Experiments and their limitations

Stated methods are used in many areas, such as marketing, health and environmental economics to study preferences. This methodology is useful because allows measurements of products not deliberated at markets, as environment goods, or to value the consequences of a policy change in welfare before its implementation or to asses market performance of hypothetical products. It is based on Thurstone`s hypothesis about human decision making made in 1927.

The nature of choice behavior modeling is rooted in the stochastic utility model expressed in equation 1. The utility of alternative *i* for the q^{th} individual can be separated in a systematic component, that can be observed and measured by the researcher, and the random component, that captures the measurement errors of the model.

$$U_{iq} = V_{iq} + \varepsilon_{iq} \tag{1}$$

Additive functions consider that total utility of the systematic term is influenced by all products' attributes. These influences are captured by the β s of equation 2, where K represents the attribute.

$$V_{iq} = \sum_{k=1}^{K} \beta_{ik} X_{ikq} \tag{2}$$

Assuming human rational behavior, individual *q* will choose the alternative *i*, among *J* alternatives, only if its utility is higher than other alternatives. More formally it is given by equation 3:

$$V_{iq} \geq V_{jq} \text{ for all } j \neq i \in A \tag{3}$$

The probability of this occurring event is:

$$\begin{aligned} P_{iq} &= \Pr ob\left(U_{iq} \geq U_{jq}, j = 1,2,...,J\right) \\ P_{iq} &= \Pr ob\left(V_{iq} + \varepsilon_{iq} \geq V_{jq} + \varepsilon_{jq}, j = 1,2,...,J\right) \\ P_{iq} &= \Pr ob\left(V_{iq} - V_{jq} \geq \varepsilon_{jq} - \varepsilon_{iq}, j = 1,2,...,J\right) \end{aligned} \tag{4}$$

Assuming that the stochastic term has a normal distribution and it is identical and independently distributed, then equation 4 can be transformed into equation 5 and derived as a Multinomial Logit Model (MNL) (equation 6) (McFadden, 1974). The log likelihood function (equation 6) is maximized using a non-linear algorithm calculating β s of equation 2 (Louviere *et al.*, 2000)

$$P_{iq} = \frac{\exp(V_{iq})}{\sum_{i=1}^{J} \exp(V_{jq})} \text{ for } i = 1,2,...,J \tag{5}$$

$$L^* = \sum_{q=1}^{Q} \sum_{j=1}^{J} f_{iq} \ln P_{iq} \tag{6}$$

The log likelihood value (L*) as well as the number of parameters (NP) are used to compare two or more competing models. Sakamoto *et al.* (1986) compared models and the best was chosen based on the Akaike Information Criteria (AIC) that is showed in equation 7. This criterion considers the lowest AIC value to select the best model.

$$AIC = -2 \times L^* + 2 \times NP \tag{7}$$

The estimated β s of the best model represents the influence of quantitative or qualitative variables on the choice process. The meaning of the β s can not be understood as partial utilities. The coefficients of effect codes mean the utility change resulting from a probability change within an attribute. As a consequence, the impact that any attribute has cannot be estimated because one of the levels is not estimated (Flynn *et al*., 2007).

Louviere *et al.* (2008), Flynn *et al*. (2007), Flynn *et al.* (2008) and Lancsar and Louviere (2008) have suggested the best-worst scaling as a solution to overcome comparisons among attributes and attributes' levels. The best-worst scaling gets parameters with the same scale and it allows those comparisons.

2.2. Properties of Best-Worst (BW) scaling and ranking theory

Best-Worst scaling, as DCE, is based on the Random Utility Theory. Finn and Louviere (1992) presented the first publication dealing with this technique, but the formal statistical and measurement properties were presented by Marley and Louviere (2005). Basically, in a best-worst choice task, respondents are asked to state the most preferable or important option and the least or less important option in a choice set. In this task respondents are not just maximizing their utilities but evaluating the maximum difference among all pairwise of options. Those models assume that subject (q) identify and calculate the difference in utility for every pair of $(U_{q,u} - U_{q,v})$ options in the choice set with J alternatives and select that pair that maximize difference in utility between them $(Y_{qj,st})$. The random utility for each ordered pair (s,t) is showed in equation 8.

$$Y_{qj,st} = U_{qj,s} - U_{qj,t} + \varepsilon_{qj,st} \quad \text{for } s,t = 1,\ldots,P \quad \text{and} \quad s \neq t \tag{8}$$

Where option s is the best option and t is the worst if $(Y_{qj,st} > Y_{qj,uv})$ for all other pairs (u,v). The relative choice probability of a given pair of options is proportional to their latent utility scale distance and, assuming that the random terms have extreme value distribution, the probability of option s and t will be chosen, respectively, as the most and the least preferred options, for subject q and choice task j, as it is showed in equation 9. Those conditions are sufficient statistics to use the Multinomial Logit Model (Marley and Louviere, 2005).

$$P_{qj}(s,t) = \frac{\exp(U_{qjs} - U_{qjt})}{\sum_{u=1}^{P} \sum_{v=1:v\neq u}^{P} \exp(U_{qju} - U_{qjv})} \quad \text{for } s \neq t \tag{9}$$

Equation 9 represents also a partially ranked choice set. That is, if there were alternatives A, B, C, D and E, being alternative A be the most preferable and E the least one, then it is possible to locate the order of the extreme alternatives A and E but not the rest. The order of the B, C, and D alternatives can only be establish in relation to A and E but not among themselves. It would be necessary to undertake a new round of best-worst decisions in order to get a complete ranking of those five alternatives. Nowadays, as in ranking and rating tasks, the best-worst ranking has been used to obtain additional information, which is obtained from the exploded process.

According to Chapman and Staelin (1986) the exploded process occurs when the probability of a ranking task is calculated. The exploded process occurs when there is a factorization of the entire choice experiment in smaller choice sets, which add new observations. Theoretically, under certain conditions, it is possible to have estimations of more efficient parameters with less variance.

The following example is given to illustrate the ranking exploding process. The preferences ordering probability of A>B>C>D>E is equal to MNL choosing A from a set {A, B, C, D, E} times the MNL choosing B from the remaining {B, C, D, E} and so on until the choice set has two alternatives, as showed in equation 10.

$$\Pr\,(\textit{ranking A, B, C, D, E}) = \frac{e^{V_A}}{\sum_{j=A,B,C,D,E} e^{V_j}} * \frac{e^{V_B}}{\sum_{j=B,C,D,E} e^{V_j}} * \frac{e^{V_C}}{\sum_{j=C,D,E} e^{V_j}} * \frac{e^{V_D}}{\sum_{j=D,E} e^{V_j}} \tag{10}$$

The exploding process is very simple but some papers assume that ranking and rating tasks induce bias respondents' behavior which violate statistical assumptions, commonly in "middle ranking" (Ben-Akiva et al., 1991; Bradley and Daly, 1994). These authors also declare that a series of "exploded" pairwise comparisons are not consistent between tasks and final analyses.

Based on Chapman and Stalin's rank logit models, Lancsar and Louviere (2008) proposed an alternative way to solve the rank and rating inconsistency. This technique improves the correspondence between the analysis of the model and the data collection. According to them, respondents first choose the best option over all alternatives of a choice set, in the example the option A, and then choose the worst option, in this case option E from the set {B, C, D, E}. The negative sign of $-V_E$ means that the indirect utility of alternative E is the worst option. The process continues to the extent that the choice set has only two alternatives, as showed in equation 11.

$$\Pr\ (\textit{best-worst ordering A, B, C, D, E}) = \frac{e^{V_A}}{\sum_{j=A,B,C,D,E} e^{V_j}} * \frac{e^{-V_E}}{\sum_{j=B,C,D,E} e^{-V_j}} * \frac{e^{V_B}}{\sum_{j=B,C,D} e^{V_j}} * \frac{e^{-V_D}}{\sum_{j=C,D} e^{-V_j}} \qquad (11)$$

Marley and Louviere (2005) stated that the rank order probability, from best to worst, must be the same than from worst to best. Thus, considering the probability sequence independency and the alternative model proposed by Lancsar and Louviere (2008) (equation 11) for an experiment with three alternatives, with preferences order A>B>C, equations (12) and (13) are presented.

$$\Pr\ (\text{from } \textit{best to worst ordering A, B, C}) = \frac{e^{V_A}}{\sum_{j=A,B,C} e^{V_j}} * \frac{e^{-V_C}}{\sum_{j=B,C} e^{-V_j}} \qquad (12)$$

$$\Pr\ (\text{from } \textit{worst to best ordering A, B, C}) = \frac{e^{-V_C}}{\sum_{j=A,B,C} e^{-V_j}} * \frac{e^{V_A}}{\sum_{j=B,C} e^{V_j}} \qquad (13)$$

So far, there has not been any work comparing respondents` abilities to choose best (maximizing utility) and worst (minimizing utility) options. Equations 12 and 13 will have worst performance than traditional DC experiments if their answers variability is different, between the best and the worst ranking options. However, it is possible to calculate a scale factor between the probabilities of choosing the best or the worst alternatives, in a choice set of tree alternatives, and the worst or best alternatives, for the remaining two alternatives to have a better model performance as it happened to Scarpa *et al.* (2009). They also studied responses' variability in a best worst task, although they analyzed the data as an exploded ranking model (equation 10). Their aim was to compare responses' variability of each best-worst statement round in large choice sets and they found that statement' variability differs in each round. In our case, we have adopted an experimental design with 3 options, as described in the next section, and in equations 12 and 13.

2.3. The experimental design

Four peaches' attributes were selected based on, the literature review about fruit quality and market tendencies, a focus group, some interviews with fruit and vegetable section managers of three retails distribution chains in Zaragoza and local market monitoring. Those attributes were the product origin, type of packing, peach size and price. Three levels were also considered for each attribute. They are listed on table 1.

Table 1. Attributes and levels employed in the experiment

Attribute	Level	Attribute	Level
Origin	From Calanda with PDO	Size	Small
	From Calanda without PDO[1]		Medium[1]
	Other places without PDO		Big
Packing	Active packing	Price	1.5 €/kg
	No active packing[1]		2.5 €/kg
	Bulk		3.5 €/kg

[1] reference levels to estimate effect codes

Price is included in the experiment as it allows Willingness to Pay (WTP) calculations for the other attributes. Price was considered as a quantitative attribute for estimation purposes and the remaining attributes as qualitative. The qualitative attributes have been estimated by code effects. Louviere *et al*. (2000) stated that effect codes are correlated in each attribute but are uncorrelated with the grand mean, unlike dummies. The effect codes should be interpreted as the difference utility in relation to a reference level.

The reference level for the origin is peaches "from Calanda without PDO". The difference of WTP between peaches "from Calanda with PDO" and "from Calanda without PDO" would be the brand value. It means how much money consumers value the guarantee of peaches with controlled quality linked to the PDO brand. The difference between WTP of peaches "from Calanda without PDO" and those "produced in other places" assesses how much consumers value the production of peaches coming from Calanda, but without the guarantees associated to the PDO brand. The expected sign of theses parameters are positive for PDO peaches and negative for peaches produced in other places.

The experiment includes two different types of packing, one normal and other active. Respondents were informed that active packing does not imply health effects and it allows keeping stocks 12 days more than with no active packing. The active packing parameter sign may be positive for some consumers who wish to store longer time. The negative sign would be expected for those consumers who, either believe that the active packing treatment has negative consequences for their health or they can also refuse packed peaches showing certain kind of neophobia.

Neophobia may contribute to positive bulk peaches' evaluations as well to the desire to touch the fruits. When consumers touch the fruit it provides them with more information about peaches' quality and they may check as well other quality cues, such as smelling, that is lost when the fruit is packed.

Different peaches' sizes were shown to respondents in the experiment. The weight of a small peach was about 160 g, a medium size was around 250 g and a big one was around 380 g. The first weight corresponds to a peach that would be refused by the PDO norms. The second is the minimum peach size accepted by the PDO norms and the largest represents a size that nobody would be able to eat at once. Normally, bigger peaches are related to higher quality, moreover there was a market segment that valued positively larger peaches up to the moment of satiating their eating capacity. Thus, the expected situation is that people have greater WTP for a medium size peach than a small one but they are not determined to select either a medium or a big size.

The configuration of the experiment corresponds to a fractional factorial design. The fractional factorial design decreases the number of combinations of a full factorial design with $\left(L^K\right)$ to $\left(L^{K-N}\right)$ combinations, which diminish the task complexity. Nine choice sets allow main estimations effects of the attributes` levels of a no label design, with 4 attributes and 3 levels in each one. According to Montgomery (2001) and Louviere *et al.* (2000) main effects explain 70 to 90% of the total variance, while two ways interaction effects only explain 5 to 15%. The estimation of two ways interaction effects would need a great number of choice sets and no references were found about analyzing this type of best-worst experiments in blockings. It was esteemed that only main effects would provide enough information.

No biased estimators are obtained if expected parameters converge to real values and efficient parameters as those that have the minimum variance. To get non biased and efficient parameters, attribute's levels were combined, following suggestions of Street *et al*. (2005). Their strategies to construct a statically efficient experiment design are based on modular mathematic, which first selects profiles from a full factorial design and then there is a generation of choice sets` options based on the first selected profiles.

Those strategies generated a balanced and orthogonal design. A design is balanced when each level of each attribute appears, in each choice set, only once. Thus, each level has the same probability to be chosen. A design is orthogonal when there is no correlation among attributes` levels. Its D-efficiency index was checked on the internet home page suggested by Pihlens *et al*. (2008), and the estimated value is 100%, so the estimated parameters are efficient and not biased.

Respondents were asked to choose the best and the worst hypothetical peaches among three alternatives in a choice set or buying situation. It allows having complete ranking alternatives in each choice set. Table 2 provides an example of a choice set. In this case, the most preferable peach would be alternative A, followed by alternatives C and B.

The questionnaire was applied to consumers attending two hypermarkets, in Zaragoza city (Spain), at the end of October 2008, when PDO Calanda peaches marketing season was finishing. Respondents spent more or less 25 minutes answering the questionnaire and they were offered, as a gift, one kilogram of peaches with PDO Calanda.

Table 2. Example of a choice set in the experiment

Least preferable	Situation 4	Most preferable
	Alternative A	
	2.5 €/kg	
☐	From Calanda with PDO	☑
	Bulk	
	Medium	
	Alternative B	
	3.5 €/kg	
☑	Other place	☐
	No active packing	
	Big	
	Alternative C	
	1.5 €/kg	
☐	From Calanda without PDO	☐
	Active packing	
	Small	

3. Results

The first question of the questionnaire asked respondents if they had consumed PDO Calanda peaches in the least two years. It was a control question and the aim was to interview respondents who somehow knew the product. At the end of the questionnaire there was a request to interview only one person per family. The intent of this warning was to avoid repetitions and over emphasizing similar profiles in order to get representative information of Zaragoza population. There were 318 valid questionnaires.

The table 3 shows the sample socio-demographic characteristics. Age, educational level, family income and professional activity are disaggregated by gender and the information is given by number of persons. The sample has a majority of women (59.1%). The age range is very broad and there are people from 18 years old to 81 years old, although females are a little beat younger because proportionally they have a greater proportion of 30 or less years old and less proportion with more than 50 years old than men. The education level is similar in both genders and the sample is composed, mostly, for elemental (30%) and medium education level (41%). Families' monthly income of 17.6% of respondents are higher than 3,000 euro, 48% of the sample is between 1,500 and 3,000 euro and for the remaining (34%) is lower than 1,500 euro. More or less 50% of de respondents have full time activity outside home and 39% full time inside, and a great percentage of the latter group is retired.

Comparisons between the socio-demographic sample information and Zaragoza city census information (INE, 2008), provides some substantial features. Thus, the sample has 7.5% more females than in the total population and 4.4% less for ages between 18 and 30 years old. Proportionally the sample has fewer respondents with medium education level than the overall population and more respondents with university studies.

The Biogeme version 1.7, a free software package to estimate discrete choice models (Barbiere, 2008), was used to analyze the data. Table 4 presents the statistical analysis. The first part of this table shows the estimated betas of each attribute level with their respective inference statistics. Almost all parameters are statistically significant at 1% confidence level. Only bulk peaches are not significant in some models.

Table 3. Respondents` socio-demographics characteristics

	Age (nº of persons)		
Gender	<=30	>30 and <=50	>50
Male	18	60	52
Female	32	87	69
	Education level (nº of persons)		
Gender	Elemental	Medium	University
Male	33	54	43
Female	54	77	57
	Income (nº of persons)		
Gender	<=1500 euros	>1500 and <=300	>3000
Male	48	57	25
Female	60	97	31
	Activity (nº of persons)		
Gender	Inside	Outside partially	Outside full
Male	45	12	73
Female	80	25	83

Source: own elaboration

As all qualitative parameters were calculated using effect codes, all reference attributes levels have a parameter value equal to the sum of the rest of the parameters linked to the same attribute. Thus the interpretation of the parameters values also need to be considering the reference levels. Thus, the probability of choosing peaches produced in the Calanda area increases if peaches have the PDO label.

Table 4. Estimated coefficients and statistical parameters

	Exploded models		**Exploded models with scale factor**		**Traditional DC experiments**	
	Best - worst	**Worst - Best**	**Best - worst**	**Worst - best**	**Best**	**Worst**
From Calanda with PDO	0.83^{**}	0.88^{**}	0.46^{**}	0.86^{**}	1.02^{**}	0.82^{**}
From Calanda without PDO$^{(1)}$	0.08	0.09	-0.01	0.10	0.02	0.02
Other places without PDO	-0.75^{**}	-0.79^{**}	-0.48^{**}	-0.77^{**}	-1.00^{**}	-0.80^{**}
Active packing	-0.15^{**}	-0.11^{**}	-0.08^{**}	-0.11^{**}	-0.21^{**}	-0.11^{**}
Normal packing$^{(1)}$	-0.17	-0.16	-0.06	-0.16	-0.17	-0.21
Bulk	$-0.02^{(NS)}$	-0.05^{*}	$0.02^{(NS)}$	-0.05^{**}	$0.04^{(NS)}$	-0.10^{*}
Price	-0.28^{**}	-0.28	-0.15	-0.28	-0.27	-0.34
Small size peach	-0.21^{**}	-0.22^{**}	-0.13^{**}	-0.22^{**}	-0.28^{**}	-0.24^{**}
Medium size peach$^{(1)}$	-0.16	-0.11	-0.09	-0.11	-0.13	-0.14
Big size peach	0.05^{**}	0.11^{**}	0.03^{**}	0.11^{**}	0.14^{**}	0.10^{**}
Scale factor	-	-	$2.20^{(NS)}$	$1.03^{(NS)}$	-	-
Final log-likelihood:	-4,224.4	-4,139.7	-4,155.4	-4,139.5	-2,293.4	-2,578.4
N. Observations	5,724	5,724	5,724	5,724	2,862	2,862
N. Parameters	10	10	11	11	10	10
Adjusted R^2	0.17	0.19	0.19	0.19	0.27	0.18
AIC	36.70	36.66	38.66	38.66	35.48	35.71

$^{(NS)}$ means no statistical significance, * significance at 5% of confidence and ** at 1%. $^{(1)}$ is the reference attribute level.
Source: own elaboration

Packing, in comparison to bulk, has the greatest impact on respondents' decisions (Table 5). Active packing does not convince respondents about its benefit of longer storage time. Consequently, the information provided to consumers that active packing had not consequences for their health and peaches' taste apparently had not positive influence on respondents' decisions.

Bulk peaches were most desired in comparison to packed peaches, although a high proportion (38.1%) of respondents agreed with the statement "I dislike peaches touched by other people" and this reaction was maintained, to some extent, even if people were wearing gloves (21.7%). The explanation about this observation is that consumers feel that peaches' quality decreases if fruits are touched by other consumers. However, it is compensated with the possibility of getting closer quality evaluation, by touching and smelling, and consequently has a better selection. Peaches' size has influenced respondents on their decisions. Large peaches are desirable and the highest difference is found between medium and big sizes.

Peaches size has influenced respondents on their choice decisions. The bigger the fruit is the more desirable is considered by consumers. All models have greater difference WTP value for big pieces with respect to medium size, which reaches an estimated value of around 0.7 €/kg. The difference WTP value between medium and small size peaches has high variability through the models.

Table 5. Willingness to Pay (WTP), in euro, to move from one level to other

	Exploded models		**Exploded models with scale factor**		**Traditional DCE**	
	Best - worst	**Worst - best**	**Best - worst**	**Worst - best**	**Best**	**Worst**
From without to with PDO	2.63	2.79	3.28	2.76	3.69	2.37
From Calanda to other place	-2.92	-3.12	-3.19	-3.10	-3.78	-2.42
From normal to active packing	0.07	0.17	-0.12	0.18	-0.15	0.29
From packing to bulk	0.54	0.40	0.52	0.41	0.77	0.32
From medium to small size	-0.18	-0.40	-0.24	-0.39	-0.53	-0.28
From medium to big size	0.73	0.78	0.87	0.77	1.02	0.70

Source: own elaboration

The model of traditional Discrete Choice (DC) experiment, when respondents pick the highest utility alternative, has fewer observations than the other models. Nevertheless, it has the lowest IAC index value showing the best fit. As a consequence of the exploding process the number of observations increases, however in this case exploded models are less efficient than traditional DC models. Best options prevail as better choosing criteria than worst options.

The application of a scale factor increases the efficiency of exploded models as detected by Scarpa *et al*. (2009). The biggest log-likelihood index change, when the scale factor is included, occurs in best to worst exploded models. This index tends to converge with the same index of the model following the sequence of worst to best, that accomplish with the statement of Marley and Louviere (2005), who established that the ordering probability from best to worst must be the same as from worst to best.

4. Final remarks

The selection of 4 attributes and 3 levels for each attribute seems reasonable to study characteristics for PDO Calanda peaches as a good way to discriminate with other peaches. The modern distribution has a tendency to use more fruits in bags and packs but Spanish consumers send clear signals that they still prefer to buy peaches on bulk. Not even so, but they are willing to pay more for bulk peaches than packed to compensate probably for damaged fruit when consumers touch the fruit. This finding is important when considering more sophisticated active packing which involves higher costs. The Spanish market still prefers big fruits conveying that probably one piece is share with other family members. However, this habit might be more difficult in the future because the number of single homes is increasing very rapidly.

Models show that consumers have a tendency to discriminate better best from worst options as they normally look around searching peaches that satisfy them the most. Probably they were very determined to choose PDO Calanda peaches over the rest and this option prevailed over other combinations of attributes and levels. The variance around the best option was quite different from the worst option and ordering was not probably a good sequence to analyze choices, so models of that nature did not perform as good as without ordering. On top, as it was a hypothetical situation, it probably induced to overstate the price they were willing to pay for PDO Calanda peaches.

Acknowledgment

This paper has been supported by INIA project, PET 2007-09_C5, financed with FEDER funds.

References

1. Barbiere, M. (2008), *Estimation of discrete choice models with BIOGEME 1.7*, biogeme.epfl.ch.
2. Ben-Akiva, M., Morikawa, T. and Shiroshi, F. (1991), "Analysis of the reliability of preference ranking data", Journal of Business Research, Vol. 23, 253–268.
3. Boletin Oficial de Aragón (1999), *Order of the Department of Agriculture, that pass the Norms of the Protected Designation of Origen "Peaches from Calanda"*, Order (BOA) Nº 114, of 1999 August 25th.
4. Bradley, M. and Daly, A. (1994), "Use of the logit scaling approach to test for rank-order and fatigue effects in state preference data", Transportation, Vol. 21, 167–184.
5. Bruhn, C.M. (1994), "Consumer and retailes satisfaction with the quality and size of California peaches and nectarines", Journal of Food Quality, Vol. 18, pp. 241–256.
6. Chapman, R.G. and Staelin, R. (1986), "Exploiting rank ordering choice set data within the stochastic utility model", Journal of Marketing Research, Vol. 19, pp. 288–301.
7. Crisosto, C.H., Crisosto, G.M. and Garner, D. (2005), "Understanding tree fruit consumer acceptance", ISHS, Acta Horticulturae, nº 682, pp. 865–870.
8. Finn, A. and Louviere, J.J. (1992), "Determining the appropriate response to evidence of public concern: the casa of food safety", Journal of Public Policy and Marketing, Vol. 50, 63–75.
9. Flynn, T.N., Louviere, J.J., Peters, T.J. and Coast, J. (2008), "Estimating preferences for dermatology consultation using Best-Worst Scaling: Comparison of various methods", BMC Medical Research Methodology, Vol. 8, nº 76.
10. Flynn, T.N., Louviere, J.J., Peters, T.J. and Coast, J. (2007), "Best-Worst Scaling: What it can do for Health care research and how to do it", Journal of Health Economics, Vol. 26, nº 1, pp. 171–189.
11. INE – National Statistic Institute of Spain (2008), Main census data of Zaragoza city in 2001, http://www.ine.es/inebmenu/mnu_cifraspob.htm
12. Lancsar, E. and Louviere, J.J. (2008), "Estimating Individual level discrete choice models and welfare measures using best worst choice experiments and sequential best worst MNL", *University of Technology, Centre for the Study of Choice (Censoc)*, working paper nº 08 – 004.
13. Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000), *Stated Choice Methods: Analysis and application*, Cambridge University Press, Cambridge.
14. Louviere, J.J., Street, D., Burgess, L., Wasi, N., Islam, T. and Marley, A.A.J. (2008), "Modelling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information", Journal of Choice Modelling, Vol. 1, nº 1, pp. 128–163.
15. McFadden, D. (1974), *Conditional logit analysis of qualitative choice behaviour*, Zarembka, P. (ed.), Frontiers in econometrics, Academic Press, New York.
16. Mainar, M.A. (2006), "Calanda en una bolsa, Surcos de Aragón", nº 99, pp. 40–43.
17. Marley, A.A.J. and Louviere, J.J. (2005), "Some probabilistic models of best, worst and best – worst choices", Journal of Mathematical Psychology, Vol. 49, pp. 464–480.
18. Montgomery, D.C. (2001), *Design and analysis of experiments*, Fifth edition, John Wile and sons, New York.
19. Pihlens, D.A., Street, D.J. and Burges, L. (2008), "Designing generic discrete choice experiments", *Proceedings of I workshop on valuation methods in agro-food and environmental economics: State-of-the art future challenges in stated choice experiments, Castelldefels3rd – 4th July*.
20. Predieri, S., Ragazzini, P. and Rondelli, R. (2006), "Sensory evaluation and peach fruit quality", ISHS, Acta Horticulturae, nº 713, pp. 429–434.
21. Polo, M. C. (2007), *La comercialización del melocotón de Denominación de Origen Calanda*, Master Thesis, IAMZ (Mediterranean Agronomic Institute of Zaragoza), Zaragoza.
22. Sakamoto, Y., Ishiguro, M. and Kitagawa (1986), *Akaike information criterion statistics*, Reidel, Dordrecht, Netherlands.

23. Scarpa, R., Notaro, S., Raffelli, R., Pihlens, D. and Louviere, J.J. (2009), "Exploring scale effects of best/worst rank order choice data to estimate visitor`s benefits from alpine transhumance", *Proceedings of International Choice Modelling Conference, Leeds, 30th March – 1st April.*

24. Street, D.J., Burguess, L. and Louvierre, J.J. (2005), "Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments", International Journal of Research in Marketing, Vol. 22, pp. 459–470.