

# Ordinal Logistic Regression for the Estimate of the Response Functions in the Conjoint Analysis

Amedeo De Luca

Dipartimento di Scienze Statistiche, Università Cattolica del Sacro Cuore di Milano, Milan, Italy.  
Email: amedeo.deluca@unicatt.it

Received June 28<sup>th</sup>, 2011; revised August 29<sup>th</sup>, 2011; accepted September 28<sup>th</sup>, 2011.

## ABSTRACT

*In the Conjoint Analysis (COA) model proposed here—a new approach to estimate more than one response function—an extension of the traditional COA, the polytomous response variable (i.e. evaluation of the overall desirability of alternative product profiles) is described by a sequence of binary variables. To link the categories of overall evaluation to the factor levels, we adopt—at the aggregate level—an ordinal logistic regression, based on a main effects experimental design. The model provides several overall desirability functions (aggregated part-worths sets), as many as the overall ordered categories are, unlike the traditional metric and non metric COA, which gives only one response function. We provide an application of the model and an interpretation of the main effects.*

**Keywords:** Aggregate Level Analysis, Conjoint Analysis, Ordinal Logistic Regression, Response Function

## 1. Introduction

Since the 1970s, Conjoint Analysis (COA) has received considerable attention as a major set of techniques for measuring buyers' trade-offs among multiattributed products and services [1].

Since that time many new developments in COA have been reported.

The purpose of this article is to give a new contribution to the problem of the conjoint measurement in order to quantify judgmental data.

The model proposed here is an extension of the traditional COA approach.

While in the traditional full-profile COA [1] the respondent expresses preferences by rating or ranking distinct product profiles, in our model we assume that the respondent evaluative judgement  $Y_k$  on the overall desirability consists in a choice of one of the  $k$  ( $k = 1, 2, \dots, K$ ) desirability categories for each of  $S$  hypothetical product profiles, chosen from a sample of respondents.

The proposed approach also differs from the "Choice-Based Conjoint" analysis (CBC) model in which the respondent expresses preferences by choosing concepts from sets of concepts (discrete choice modelling).

In the proposed approach the ordinal response variable is described by an ordered logit model, that directly incorporates the order of the categories of the  $Y_k$ .

To link the categories of overall evaluation to the factor levels, we adopt a cumulative logit model [2] at the aggregate level (*pooled model*) [3].

The main novelty value in our approach—which is a further development—is that one set of aggregated part-worths is estimated in connection with each category  $Y_k$ , as many as the overall ordered categories are ( $K$ ), unlike the traditional metric and non metric COA and CBC analysis, which give only one set of aggregated part-worths (response function).

Thus with our approach it is possible to make a cross-check of the effects of the attribute levels on the different  $k$  categories of  $Y_k$ .

Moreover, the proposed model provides the following advantages: the use of the probability  $P_{ks}$  as an average response, which does not require preliminary scale adjustments to render the preference scale "metric" (in the non metric COA) and a cross-check of the effects of the attribute levels on the different categories of the overall. This allows us to verify the basic coherence of the results of the model, unlike the classical approach adopted in literature.

This paper is structured as follows: after the presentation of the proposed model and the estimation method of response functions (the cumulative logit model), a concrete application of the approach follows; then the meaning of the proposed interpretative model and some empirical results are described, to conclude with a

statement of the benefits offered by the proposed model—compared to the approaches known in literature—as far as the methodology of the COA is concerned.

## 2. Estimation of Response Functions in the Conjoint Analysis: The Cumulative Logit Model

The *cumulative* logit model proposed here (that directly incorporates the order of the categories of the *overall* desirability of alternative concepts of the product) concerns the full-profile coa. It is based on overall desirability categories chosen by a sample of respondents, for each of  $S$  hypothetical product profiles.

The number of profiles  $S$ , resulting from the total number of possible combinations of levels of the  $M$  attributes or factors ( $X$ ) of a product, constitute a *full-factorial* experimental design.

The focus of our paper is to estimate the relationship between dependent and independent variables.

It is assumed that the global or *overall* evaluation (polytomous dependent ordinal variable  $Y$ ) of a product consists in the choice of one of the ordered categories  $k = 1, 2, \dots, K$  (in our application  $K = 5$ ) on scale 1 - 5 (1 = “less desirable”, 5 = “most desirable”).

In terms of probabilities, the effects of the factors express the variations of the probabilities  $P_{ks}$ —if  $k$  is the overall category—associated with the vector  $\tilde{\mathbf{z}}_s$  corresponding to the combination  $s$  ( $s = 1, 2, \dots, S$ ) of levels of the  $M$  factor, as follows:

$$P(Y_k = 1 | \tilde{\mathbf{z}}_s) = \frac{\exp(\delta_k + \tilde{\delta}' \tilde{\mathbf{z}}_s)}{1 + \exp(\delta_k + \tilde{\delta}' \tilde{\mathbf{z}}_s)} \tag{1}$$

$$= F_k(\tilde{\mathbf{z}}_s)$$

where

$\delta_k$  is the intercept term in regression;

$\tilde{\delta}$  is the unknown vector of regression coefficients of the factors;

$\tilde{\mathbf{z}}_s$  is the vector of the indicator explanatory variables relative to the combination or profile  $s$ ;

$F_k(\tilde{\mathbf{z}}_s) = P(Y \leq k | \tilde{\mathbf{z}}_s)$  is the cumulative probability for response category  $k$ , when the explanatory variables take the value  $\tilde{\mathbf{z}}_s$ .

When we have a simple random sample of  $J$  respondents, for the formula (1), the sample likelihood turns out to be:

$$L(\delta) = \prod_{j=1}^J [F(\delta' \tilde{\mathbf{z}}_j)]^{y_j} [1 - F(\delta' \tilde{\mathbf{z}}_j)]^{1-y_j} \tag{2}$$

where

$\tilde{\mathbf{z}}_j, j = 1, 2, \dots, J$ , summarizes the underlying conditions related to the generic  $j$ -th respondent of the sam-

ple;

$y_j = 1$  if the value of the dichotomized overall evaluation is  $Y_j = 1$ . This is obtained by placing in the first class all the evaluations with *desirability* judgement of class greater than or equal to a given category  $k$  and placing in the other the remaining ones (ordinal logistic regression).

It is shown that, under suitable conditions, regarding the behaviour of the arrays  $\tilde{\mathbf{z}}_j$  as long as  $J$  increases, if the system of likelihood  $\partial \ln L / \partial \delta_k = 0, k = 1, 2, \dots, K$ , has a solution, this is of absolute maximum, and defines a consistent estimator  $\tilde{\delta}$  of the parametric vector  $\underline{\delta}$ , asymptotically normally distributed. See, for example, [4].

To estimate said probabilities  $P(Y_k = 1 | \tilde{\mathbf{z}}_s)$  we use an *aggregate* level model across the  $J$  homogeneous research respondents [5], whose evaluations, on each product profile, are considered  $J$  repeated observations.

At this point it is necessary to estimate the relationship between  $Y_k$  ( $k = 1, 2, \dots, K$ ) dependent variable (*overall* judgment category) and  $m = 1, 2, \dots, M$  (in our application  $M = 3$ ) qualitative independent variables (product attributes or factors  $X$ ), with levels  $l = 1, 2, \dots, l_m$  (in our application:  $l_1 = 3; l_2 = 2, l_3 = 3$ ).

The  $K$  overall categories ( $Y_k$ ) are codified as  $K$  indicator variables. Also the independent variables are codified as  $Z$  indicator variables (for each variable we have defined a set of 0-1 indicator variables  $Z_l^{(m)}, l = 1, 2, \dots, l_m$ , so that—for one  $m$  factor— $Z_l^{(m)} = 1$  if category  $l$ th is observed, while in the all other cases  $Z_l^{(m)} = 0$ ).

To obtain univocal estimates of the parameters, the column concerning the first variable of each set has been dropped. Therefore the  $k$ th cumulative response probability is:

$$P_{ks}((Y \leq k | \tilde{\mathbf{z}}_s)) = F_k(\tilde{\mathbf{z}}_s)$$

$$= \pi_1(\tilde{\mathbf{z}}_s) + \pi_2(\tilde{\mathbf{z}}_s) + \dots + \pi_k(\tilde{\mathbf{z}}_s),$$

$$k = 1, 2, \dots, K, \tag{2}$$

where  $\pi_k(\tilde{\mathbf{z}}_s)$  is the probability of the response  $k$  associated with the reduced vector  $\tilde{\mathbf{z}}_s = [1, z_{12}, z_{13}, z_{22}, z_{32}, z_{33}]$  of the explicative variables that indicate the assessment values.

The cumulative probabilities reflect the ordering, with:

$$P_{ks}(Y \leq 1 | \tilde{\mathbf{z}}_s) \leq P_{ks}(Y \leq 2 | \tilde{\mathbf{z}}_s) \dots \leq (P_{ks} \leq K | \tilde{\mathbf{z}}_s)$$

where  $P_{ks}(Y \leq K | \tilde{\mathbf{z}}_s) = 1$ .

The model for cumulative probabilities does not use the final cumulative probability  $P_{ks}(Y \leq K | \tilde{\mathbf{z}}_s)$ , since it is necessarily equal to 1.

In the configured model, owing to the interrelationship between the  $K$  dependent variables, the  $K$ th equation can be drawn from the remaining  $q = K - 1$  equations.

The cumulative logits of the first  $(K - 1)$  cumulative probabilities are:

$$L_k(\tilde{\mathbf{z}}_s) = \text{logit} \left[ F_k(\tilde{\mathbf{z}}_s) \right] = \ln \left[ \frac{F_k(\tilde{\mathbf{z}}_s)}{1 - F_k(\tilde{\mathbf{z}}_s)} \right]$$

$$= \ln \left[ \frac{\pi_1(\tilde{\mathbf{z}}_s) + \pi_2(\tilde{\mathbf{z}}_s) + \dots + \pi_k(\tilde{\mathbf{z}}_s)}{\pi_{k+1}(\tilde{\mathbf{z}}_s) + \pi_{k+2}(\tilde{\mathbf{z}}_s) + \dots + \pi_K(\tilde{\mathbf{z}}_s)} \right] \quad (3)$$

$$= \delta_k + \underline{\tilde{\delta}} \tilde{\mathbf{z}}$$

with  $k = 1, 2, \dots, K - 1$ ;  
where

$\tilde{\mathbf{z}}_s$  is the vector of the reduced matrix  $\tilde{\mathbf{Z}}$  (in which it has been dropped out one of the indicator variables for each level of the  $M$  factors  $X$ );

$\delta_k$  is the constant term associated to the reference category (the  $\delta_k$  are called *cutpoint* parameters. They are not decreasing in  $k$ , since the cumulative logit is an increasing function of  $F_k(\tilde{\mathbf{z}}_s)$ , which is itself increasing in  $k$  for fixed  $\tilde{\mathbf{z}}_s$ );

$\underline{\tilde{\delta}}$  is the vector of the unknown coefficients.

In Equation (3)  $\underline{\tilde{\delta}}$  does not have a  $k$  subscript (*Proportional Odd Assumption*—POA; see [6]), so the model assumes the same effects as  $\tilde{\mathbf{Z}}$  for all  $K - 1$  on all cumulative logit results, in a parsimonious model for ordinal data.

When this model fits well it requires a single parameter rather than  $K - 1$  parameters to describe the effect of  $\tilde{\mathbf{z}}$ .

### 3. Application of the Cumulative Logit Model

The model was applied to the overall desirability evaluations expressed on the  $K = 5$  categories by a sample of  $J = 79$  users on  $S = 18$  new profiles of mobile phones.

The  $M = 3$  experimental factors and levels where:  $X_1 =$  “weight” (with levels:  $\leq 94$  grams,  $95 - 105$  grams,  $>105$  grams);  $X_2 =$  “autonomy” ( $\leq 200$  h,  $>200$  h);  $X_3 =$  “price” ( $<200$  €,  $200 - 300$  €,  $>300$  €).

The  $K$  overall categories ( $Y_k$ ) are codified as  $K$  indicator variables (Table 1); also the independent variables are codified as indicator variables ( $Z$ ) (Table 2).

**Table 1. Disjunctive binary coding of overall evaluations ( $Y_k$ ) categories.**

Overall evaluation ( $Y_k$ )	Indicator variables				
	$Y_1$	$Y_2$	$Y_3$	$Y_4$	$Y_5$
$k = 1$	1	0	0	0	0
$k = 2$	0	1	0	0	0
$k = 3$	0	0	1	0	0
$k = 4$	0	0	0	1	0
$k = 5$	0	0	0	0	1

**Table 2. Disjunctive binary coding of factors: “weight”, “autonomy”, “price”.**

Predictor variables and levels ( $I$ )	Indicator variables		
<b>Weight</b>	$Z_1^{(1)}$	$Z_2^{(1)}$	$Z_3^{(1)}$
$\leq 94$ grams	1	0	0
$95 - 105$ grams	0	1	0
$>105$ grams	0	0	1
<b>Predictor variables and levels (<math>I</math>)</b>	<b>Indicator variables</b>		
<b>Autonomy</b>	$Z_1^{(2)}$	$Z_2^{(2)}$	
$\leq 200$ h	1	0	
$>200$ h	0	1	
<b>Predictor variables and levels (<math>I</math>)</b>	<b>Indicator variables</b>		
<b>Price</b>	$Z_1^{(3)}$	$Z_2^{(3)}$	$Z_3^{(3)}$
$<200$ €	1	0	0
$200 - 300$ €	0	1	0
$>300$ €	0	0	1

The reduced matrix  $\tilde{\mathbf{Z}}$  of the indicator variables of the experimental design is as follows (see § 2):

$$\tilde{\mathbf{Z}} = \begin{bmatrix} 1 & | & 0 & 0 & | & 0 & | & 0 & 0 \\ 1 & | & 0 & 0 & | & 0 & | & 1 & 0 \\ 1 & | & 0 & 0 & | & 0 & | & 0 & 1 \\ 1 & | & 0 & 0 & | & 1 & | & 0 & 0 \\ 1 & | & 0 & 0 & | & 1 & | & 1 & 0 \\ 1 & | & 0 & 0 & | & 1 & | & 0 & 1 \\ 1 & | & 1 & 0 & | & 0 & | & 0 & 0 \\ 1 & | & 1 & 0 & | & 0 & | & 1 & 0 \\ 1 & | & 1 & 0 & | & 1 & | & 1 & 0 \\ 1 & | & 1 & 0 & | & 1 & | & 0 & 1 \\ 1 & | & 1 & 0 & | & 1 & | & 0 & 0 \\ 1 & | & 0 & 1 & | & 0 & | & 0 & 0 \\ 1 & | & 0 & 1 & | & 0 & | & 1 & 0 \\ 1 & | & 0 & 1 & | & 0 & | & 0 & 1 \\ 1 & | & 0 & 1 & | & 1 & | & 0 & 0 \\ 1 & | & 0 & 1 & | & 1 & | & 1 & 0 \\ 1 & | & 0 & 1 & | & 1 & | & 0 & 1 \end{bmatrix}$$

The application was made using the *Proportional Option Assumption—POA*; the model is estimated with a PLUM-Ordinal regression procedure, available in SPSS 10 and the following ones.

The parameters were estimated using the maximum likelihood method linked to the POA hypothesis (the *Fisher's Scoring* optimisation algorithm, [7]).

The judgement evaluations are pooled across respondents (*pooled model*) and the novelty value in our approach is that one set of aggregated part-worths is estimated in connection with each overall category  $Y_k$  (see **Table 3**).

#### 4. Meaning of the Proposed Interpretative Model and Some Empirical Results

Model (3) can be expressed as follows (the coefficients were estimated as shown in the §2, with reference to (2)):

$$\begin{aligned} \tilde{L}_1(\tilde{\mathbf{z}}_s) &= \text{logit}[F_1(\tilde{\mathbf{z}})] = \ln \left[ \frac{F_1(\tilde{\mathbf{z}}_s)}{1 - F_1(\tilde{\mathbf{z}}_s)} \right] \\ &= \ln \left[ \frac{\tilde{\pi}_1}{\tilde{\pi}_2 + \tilde{\pi}_3 + \tilde{\pi}_4 + \tilde{\pi}_5} \right] \\ &= \tilde{c}_1 + \tilde{\delta}_2^1 \tilde{z}_{12} + \tilde{\delta}_3^1 \tilde{z}_{13} + \tilde{\delta}_2^2 \tilde{z}_{22} \\ &\quad + \tilde{\delta}_2^3 \tilde{z}_{32} + \tilde{\delta}_3^3 \tilde{z}_{33} \\ &= -4.409 - 0.944\tilde{z}_{12} - 1.93\tilde{z}_{13} \\ &\quad + 1.041\tilde{z}_{22} - 1.197\tilde{z}_{32} - 2.355\tilde{z}_{33} ; \end{aligned}$$

**Table 3. Estimates of four set of the aggregated part-worths utilities of the COA model ordinal logistic regression.**

	Equations	Estimated coefficient	Standard error	df	Wald $\chi^2$	p-value
Intercept	$Y = 1$	-4.409	0.171	1	662.647	0.000
	$Y = 2$	-2.600	0.144	1	327.567	0.000
	$Y = 3$	-0.610	0.126	1	23.470	0.000
	$Y = 4$	1.512	0.140	1	116.584	0.000
Factor	Levels					
Weight	$z_{12}$	-0.944	0.122	1	59.439	0.000
	$z_{13}$	-1.930	0.129	1	223.215	0.000
Autonomy	$z_{22}$	1.041	0.101	1	105.829	0.000
Price	$z_{32}$	-1.197	0.124	1	93.015	0.000
	$z_{33}$	-2.355	0.134	1	310.179	0.000

$$\begin{aligned} \tilde{L}_2(\tilde{\mathbf{z}}_s) &= \text{logit}[F_2(\tilde{\mathbf{z}})] = \ln \left[ \frac{F_2(\tilde{\mathbf{z}}_s)}{1 - F_2(\tilde{\mathbf{z}}_s)} \right] \\ &= \ln \left[ \frac{\tilde{\pi}_1 + \tilde{\pi}_2}{\tilde{\pi}_3 + \tilde{\pi}_4 + \tilde{\pi}_5} \right] \\ &= \tilde{c}_2 + \tilde{\delta}_2^1 \tilde{z}_{12} + \tilde{\delta}_3^1 \tilde{z}_{13} + \tilde{\delta}_2^2 \tilde{z}_{22} \\ &\quad + \tilde{\delta}_2^3 \tilde{z}_{32} + \tilde{\delta}_3^3 \tilde{z}_{33} \\ &= -2.6 - 0.944\tilde{z}_{12} - 1.93\tilde{z}_{13} \\ &\quad + 1.041\tilde{z}_{22} - 1.197\tilde{z}_{32} - 2.355\tilde{z}_{33} ; \end{aligned}$$

$$\begin{aligned} \tilde{L}_3(\tilde{\mathbf{z}}_s) &= \text{logit}[F_3(\tilde{\mathbf{z}})] = \ln \left[ \frac{F_3(\tilde{\mathbf{z}}_s)}{1 - F_3(\tilde{\mathbf{z}}_s)} \right] \\ &= \ln \left[ \frac{\tilde{\pi}_1 + \tilde{\pi}_2 + \tilde{\pi}_3}{\tilde{\pi}_4 + \tilde{\pi}_5} \right] \\ &= \tilde{c}_3 + \tilde{\delta}_2^1 \tilde{z}_{12} + \tilde{\delta}_3^1 \tilde{z}_{13} + \tilde{\delta}_2^2 \tilde{z}_{22} \\ &\quad + \tilde{\delta}_2^3 \tilde{z}_{32} + \tilde{\delta}_3^3 \tilde{z}_{33} \\ &= -0.61 - 0.944\tilde{z}_{12} - 1.93\tilde{z}_{13} \\ &\quad + 1.041\tilde{z}_{22} - 1.197\tilde{z}_{32} - 2.355\tilde{z}_{33} ; \end{aligned}$$

$$\begin{aligned} \tilde{L}_4(\tilde{\mathbf{z}}_s) &= \text{logit}[F_4(\tilde{\mathbf{z}})] = \ln \left[ \frac{F_4(\tilde{\mathbf{z}}_s)}{1 - F_4(\tilde{\mathbf{z}}_s)} \right] \\ &= \ln \left[ \frac{\tilde{\pi}_1 + \tilde{\pi}_2 + \tilde{\pi}_3 + \tilde{\pi}_4}{\tilde{\pi}_5} \right] \\ &= \tilde{c}_4 + \tilde{\delta}_2^1 \tilde{z}_{12} + \tilde{\delta}_3^1 \tilde{z}_{13} + \tilde{\delta}_2^2 \tilde{z}_{22} \\ &\quad + \tilde{\delta}_2^3 \tilde{z}_{32} + \tilde{\delta}_3^3 \tilde{z}_{33} \\ &= 1.512 - 0.944\tilde{z}_{12} - 1.93\tilde{z}_{13} \\ &\quad + 1.041\tilde{z}_{22} - 1.197\tilde{z}_{32} - 2.355\tilde{z}_{33} \end{aligned}$$

Out of a reading of the coefficients (*effects*) in **Table 3**, we can see the modalities of the factors that contribute to the increase/decrease the  $\hat{p}_k$  values ( $k = 1, 2, 3, 4, 5$ ) and, consequently, the relative importance of each attribute as well as which levels of each attribute are most preferred.

**Table 3** points out that the intercept value related to the fourth equation, associated to the global evaluation  $Y_4$ , is of opposite compared to the algebraic sign of the first, second and third equation (associated, respectively, to the 1, 2 and 3 judgement categories).

This allow us to verify the basic coherence of the results of the model (at least with regard to the main effects).

The factor that influences the logit more is Price, followed by Weight; less important is Autonomy.

The probabilities relative to the five overall categories are expressed, respectively, as follows:

$$\tilde{\pi}_1 = \frac{e^{-4,409+0,944\tilde{z}_{12}+1,93\tilde{z}_{13}-1,041\tilde{z}_{22}+1,197\tilde{z}_{32}+2,355\tilde{z}_{33}}}{1 + e^{-4,409+0,944\tilde{z}_{12}+1,93\tilde{z}_{13}-1,041\tilde{z}_{22}+1,197\tilde{z}_{32}+2,355\tilde{z}_{33}}}$$

$$\tilde{\pi}_2 = \frac{e^{-2,6+0,944\tilde{z}_{12}+1,93\tilde{z}_{13}-1,041\tilde{z}_{22}+1,197\tilde{z}_{32}+2,355\tilde{z}_{33}}}{1+e^{-2,6+0,944\tilde{z}_{12}+1,93\tilde{z}_{13}-1,041\tilde{z}_{22}+1,197\tilde{z}_{32}+2,355\tilde{z}_{33}}} - \tilde{\pi}_1$$

$$\tilde{\pi}_3 = \frac{e^{-0,61+0,944\tilde{z}_{12}+1,93\tilde{z}_{13}-1,041\tilde{z}_{22}+1,197\tilde{z}_{32}+2,355\tilde{z}_{33}}}{1+e^{-0,61+0,944\tilde{z}_{12}+1,93\tilde{z}_{13}-1,041\tilde{z}_{22}+1,197\tilde{z}_{32}+2,355\tilde{z}_{33}}} - \tilde{\pi}_1 - \tilde{\pi}_2$$

$$\tilde{\pi}_4 = \frac{e^{1,512+0,944\tilde{z}_{12}+1,93\tilde{z}_{13}-1,041\tilde{z}_{22}+1,197\tilde{z}_{32}+2,355\tilde{z}_{33}}}{1+e^{1,512+0,944\tilde{z}_{12}+1,93\tilde{z}_{13}-1,041\tilde{z}_{22}+1,197\tilde{z}_{32}+2,355\tilde{z}_{33}}} - \tilde{\pi}_1 - \tilde{\pi}_2 - \tilde{\pi}_3$$

$$\tilde{\pi}_5 = 1 - \tilde{\pi}_1 - \tilde{\pi}_2 - \tilde{\pi}_3 - \tilde{\pi}_4$$

In order to empirically assess the predictive capacity of the estimated model, **Table 4** shows the probabilities estimated for all the modality combinations (experimental conditions “s”) of the explanatory variables and the corresponding values of the observed proportions.

We notice a satisfactory model fitting, as the predicted probabilities turn out to be very near the corresponding

proportions for all the modality combinations of the experimental design.

**Table 5** supplies, for each product profile “s”, the odds value relative to the four probability functions  $\tilde{\pi}_k$ , that is, the values:

$$odds1 = \left[ \frac{\tilde{\pi}_1}{\tilde{\pi}_2 + \tilde{\pi}_3 + \tilde{\pi}_4 + \tilde{\pi}_5} \right] = e^{\delta_1 + \delta'z}$$

$$odds2 = \left[ \frac{\tilde{\pi}_1 + \tilde{\pi}_2}{\tilde{\pi}_3 + \tilde{\pi}_4 + \tilde{\pi}_5} \right] = e^{\delta_2 + \delta'z}$$

$$odds3 = \left[ \frac{\tilde{\pi}_1 + \tilde{\pi}_2 + \tilde{\pi}_3}{\tilde{\pi}_4 + \tilde{\pi}_5} \right] = e^{\delta_3 + \delta'z}$$

$$odds4 = \left[ \frac{\tilde{\pi}_1 + \tilde{\pi}_2 + \tilde{\pi}_3 + \tilde{\pi}_4}{\tilde{\pi}_5} \right] = e^{\delta_4 + \delta'z}$$

**Table 4. Comparison of the predicted probabilities, estimated by the COA model, and the corresponding proportions for all the modality combinations of the experimental design.**

s	Y <sub>1</sub>		Y <sub>2</sub>		Y <sub>3</sub>		Y <sub>4</sub>		Y <sub>5</sub>	
	Probability	Proportion	Probability	Proportion	Probability	Proportion	Probability	Proportion	Probability	Proportion
	$\tilde{\pi}_1(\mathbf{z}_s)$	$p_1$	$\tilde{\pi}_2(\mathbf{z}_s)$	$p_2$	$\tilde{\pi}_3(\mathbf{z}_s)$	$p_3$	$\tilde{\pi}_4(\mathbf{z}_s)$	$p_4$	$\tilde{\pi}_5(\mathbf{z}_s)$	$p_5$
1	0.01	0.01	0.06	0.04	0.28	0.30	0.47	0.46	0.18	0.19
2	0.04	0.04	0.16	0.20	0.45	0.41	0.29	0.30	0.06	0.05
3	0.11	0.13	0.33	0.29	0.41	0.46	0.13	0.11	0.02	0.01
4	0.00	0.00	0.02	0.01	0.14	0.14	0.45	0.44	0.38	0.41
5	0.01	0.03	0.07	0.06	0.31	0.25	0.45	0.49	0.16	0.16
6	0.04	0.06	0.17	0.15	0.45	0.49	0.28	0.24	0.06	0.05
7	0.03	0.01	0.13	0.16	0.42	0.39	0.34	0.35	0.08	0.08
8	0.09	0.08	0.29	0.28	0.43	0.49	0.15	0.14	0.03	0.01
9	0.25	0.24	0.42	0.39	0.27	0.30	0.06	0.06	0.01	0.00
10	0.01	0.00	0.05	0.06	0.27	0.24	0.47	0.51	0.20	0.19
11	0.04	0.04	0.15	0.14	0.44	0.47	0.31	0.33	0.07	0.03
12	0.10	0.13	0.31	0.33	0.42	0.35	0.14	0.19	0.02	0.00
13	0.03	0.03	0.12	0.14	0.42	0.44	0.35	0.29	0.08	0.10
14	0.09	0.04	0.29	0.33	0.44	0.46	0.16	0.14	0.03	0.04
15	0.24	0.22	0.42	0.44	0.28	0.24	0.06	0.09	0.01	0.01
16	0.08	0.08	0.26	0.29	0.45	0.44	0.18	0.13	0.03	0.06
17	0.22	0.23	0.41	0.41	0.30	0.28	0.07	0.06	0.01	0.03
18	0.47	0.49	0.37	0.29	0.13	0.18	0.02	0.03	0.00	0.01

**Table 5. The odds of the 18 combinations of the experimental design.**

$s$	$c$	$z_{12}$	$z_{13}$	$z_{22}$	$z_{32}$	$z_{33}$	$L_1$	$odds1$	$L_2$	$odds2$	$L_3$	$odds3$	$L_4$	$odds4$
1	1	1	1	1	1	1	-4.409	0.012	-2.6	0.074	-0.61	0.543	1.512	4.536
2	1	1	1	1	0	1	-3.212	0.04	-1.403	0.246	0.587	1.799	2.709	15.014
3	1	1	1	1	1	0	-2.054	0.128	-0.245	0.783	1.745	5.726	3.867	47.799
4	1	1	1	0	1	1	-5.45	0.004	-3.641	0.026	-1.651	0.192	0.471	1.602
5	1	1	1	0	0	1	-4.253	0.014	-2.444	0.087	-0.454	0.635	1.668	5.302
6	1	1	1	0	1	0	-3.095	0.045	-1.286	0.276	0.704	2.022	2.826	16.878
7	1	0	1	1	1	1	-3.465	0.031	-1.656	0.191	0.334	1.397	2.456	11.659
8	1	0	1	1	0	1	-2.268	0.104	-0.459	0.632	1.531	4.623	3.653	38.59
9	1	0	1	1	1	0	-1.11	0.329	0.699	2.012	2.689	14.72	4.811	122.85
10	1	0	1	0	1	1	-4.506	0.011	-2.697	0.067	-0.707	0.493	1.415	4.116
11	1	0	1	0	0	1	-3.309	0.037	-1.5	0.223	0.49	1.632	2.612	13.626
12	1	0	1	0	1	0	-2.151	0.116	-0.342	0.71	1.648	5.197	3.77	43.38
13	1	1	0	0	1	1	-3.52	0.029	-1.711	0.181	0.279	1.322	2.401	11.034
14	1	1	0	0	0	1	-2.323	0.098	-0.514	0.598	1.476	4.375	3.598	36.525
15	1	1	0	0	1	0	-1.165	0.312	0.644	1.904	2.634	13.93	4.756	116.28
16	1	1	0	1	1	1	-2.479	0.084	-0.67	0.512	1.32	3.743	3.442	31.249
17	1	1	0	1	0	1	-1.282	0.277	0.527	1.694	2.517	12.39	4.639	103.44
18	1	1	0	1	1	0	-0.124	0.883	1.685	5.392	3.675	39.45	5.797	329.31

The odds1 of the profile 4 (*the best*: weight  $\leq 94$  g; autonomy  $> 200$  h, price  $\leq 200$  €) it is much lower (0,004) and indicates that the probability to assign 1 is 0,004 times the probability to give another response (Table 4). The probability to assign 1 or 2 is 0.026 times the probability to assign 3, 4, 5.

The odds1 of the profile 18 (*the worst*: weight  $> 105$  g; autonomy  $\leq 200$  h, price  $> 300$  €) it is much higher (0.883) and indicates that the probability to assign 1 is 0.883 times the probability to give an other response.

The probability to assign 1 or 2 exceeds 5.392 times the probability to assign 3, 4, 5.

## 5. Final Remarks

Besides these positive features, the model here proposed provides the following remarkable advantages:

1) the use of the probability  $\hat{P}_{ks}$  as an average response, which does not require scale adjustments to ren-

der the preference scale “metric”;

2) the estimate of one set of aggregated part-worths in connection with *each overall* category  $k$ ;

3) a cross-check of the effects of the attribute levels on the different  $k$  categories of  $Y_k$ ; this allows us verify the basic coherence of the results of the model;

4) the proposed model, at the aggregate level, offers the prospect of more accurate estimation, unlike the traditional conjoint methods which estimate part-worth utilities at an individual level.

Moreover, we can argue that an aggregate analysis permits the estimation of subtle interaction effects due to its ability to leverage a great deal of data across respondents.

## REFERENCES

- [1] P. E. Green and V. R. Rao, “Conjoint Measurement for Quantifying Judgmental Data,” *Journal of Marketing Re-*

- search*, Vol. 8, No. 3, 1971, pp. 355-363.
- [2] A. De Luca, "A Logit Model with a Variable Response and Predictors on an Ordinal Scale to Measure Customer Satisfaction," *Quality and Reliability Engineering International*, Vol. 22, No. 5, 2006, pp. 591-602.  
[doi:10.1002/qre.764](https://doi.org/10.1002/qre.764)
- [3] W. L. Moore, "Levels of Aggregation in Conjoint Analysis: An Empirical Comparison," *Journal of Marketing Research*, Vol. 17, No. 4, 1980, pp. 516-523.  
[doi:10.2307/3150504](https://doi.org/10.2307/3150504)
- [4] T. Amemiya, "Advanced Econometrics," Blackwell Ltd., Oxford, 1985.
- [5] A. De Luca and S. Ciapparelli, "Multivariate Logistic Regression for the Estimate of Response Functions in the Conjoint Analysis," *Proceedings MTISD 2008—Methods, Models and Information Technologies for Decision Support Systems*, Università del Salento, Lecce, 2008, pp. 23-24.
- [6] B. Peterson and F. E. Harrell, "Partial Proportional Odds Models for Ordinal Response Variables," *Applied Statistics*, Vol. 39, No. 2, 1990, pp. 205-207.  
[doi:10.2307/2347760](https://doi.org/10.2307/2347760)
- [7] P. McCullagh, "Regression Models for Ordinal Data," *Journal of Royal Statistical Society*, Vol. 42, No. 2, 1980, pp. 109-142.