DISA
Via Inama, 5
38100 Trento (Italy)

Shortlisting by Incomplete Descriptions: The Power of Combination

Luca Pedrotti   Silvia Rensi   Enrico Zaninotto

UNIVERSITÀ DEGLI STUDI
DI TRENTO

QUADERNO 113 giugno 2006

Shortlisting by Incomplete Descriptions: The Power of Combination*

Luca Pedrotti
Department of Computer and Management Science, University of Trento

Silvia Rensi
Department of Mathematics, University of Trento

Enrico Zaninotto
Department of Computer and Management Science, University of Trento

Abstract

In this paper we make use of a particular technique of data analysis to empirically study the effect of joint attributes presentation in a multi-step process of choice, in which consumers first use a simple method for shortlisting, then proceed with a closer inspection of a restrict number of alternatives. Shortlisting is based on an incomplete description of the attributes of an alternative. We focus, in particular, on the presentation couples of attributes. The mathematical framework we used is the generalized spectral analysis. We tested this method on data collected through an ad hoc survey. Thanks to this powerful machinery we were able to identify the attraction single attributes have, from the effect of their combination. The use of generalized spectral analysis to decompose data on preferences is totally new. The decomposition allows us to underline two effects: the first and second order effect. The first order effect measures the average attraction that a single feature has when it is coupled with a second one. The second order effect detects the positive (or negative) power of combination of two coupled attributes. We present here a particular case, the choice of a car, among the ones we studied, to show how the method can be used, and its power. A particular emphasis will be given to gender differences in the evaluation of car attributes in the choice process.

* We thank Ingeborg Weger for the great help she gave us to collect the interviews

1 Introduction

When choosing complex goods or services, like cars or holidays, consumers often shortlist a subset of alternatives on the basis of few attributes, highlighted by advertising or business communication. Then their choice proceeds with finer methods.

Choice methods by attributes have been widely studied: their properties have been underlined by Payne, Bettman, and Johnson (1993) and Payne, Bettman, and Johnson (1988), who give emphasis to the trade off between accuracy and effort in evaluating attributes. In this paper we empirically study the effect of an evaluation process by attributes as part of a decision process in which consumers use first a simple method for shortlisting and then proceed with a closer inspection of a restrict number of alternatives. In particular, we propose an empirical method to isolate the effect of a joint presentation of attributes, and we test the method on data collected through an ad hoc survey.

The paper is organized as follows. First, we present a model for shortlisting, which rests on two ideas. On the one hand, goods or services enter in the evaluation process as incomplete descriptions, i.e. as a subset of described characteristics in binary form. For instance, a car in a given range of prices can be described as being comfortable and fast, or elegant and safe, and so on. On the other hand, consumers use in their first step of decision a simple procedure: they have a simplified preference vector over attributes, by which they give weight only to the most preferred items. Finally, they shortlist only goods whose incomplete description matches preferences over attributes.

Secondly, we carried a survey to simulate a simple process of choice among different combination of attributes. We started with a preliminary survey (40 open interviews) to isolate the attributes mostly considered when choosing 5 goods or services (a car, a mobile phone, a desktop PC, a holiday, a restaurant). We isolated 5 attributes for each good or service. Then we made a large survey (more than 1000 interviews), asking people what had been the attribute(s) used to select a good, using both single attributes and any possible combination of two. For instance, in the case of a car, we asked if they had to choose a car in a given range of price, they would look at comfort and safe, comfort and look, and so on.

Then, making use of generalized spectral analysis see Diaconis (1988), Diaconis (1989), Diaconis and Rockmore (1990), we were able to identify the attraction single attributes have, from the effect of their combination. The use of generalized spectral analysis to decompose preferences data is totally new. Lawson and Orrison (2002) and Lawson, Orrison, and Uminsky (2003) used it to detect hidden coalitions in the vote of nine judges of the United States Supreme Court, while the nearer application was the one proposed by Rensi and Zaninotto (2004) to study preferences for parties. Being a non-model method for data analysis, its power depends on the interpretation of the decomposition we obtained. In particular first, through the analysis of first order effect, we were able to measure (in the two attributes choice) the attraction effect a single feature has, through the attributes to which it is combined. Second, and most important, through the analysis of second order effects, we were able to identify the positive (and negative) power the combination of attributes has in driving potential consumers choice. Positive second order effect can be interpreted as a sort of "cognitive complementarity".

2 The shortlisting process: concepts and definitions

Complex goods are characterized by a wide set of characteristics, making a comparison cognitively difficult. It is still under debate whether, facing a difficult choice, consumers try to single out characteristics in order to make some sort of multi-attribute choice, or try to build an overall view of the alternatives. In this paper we assume that decision makers compare the alternatives along different attributes. Nevertheless, due to the cognitive difficulty of a multi-attribute comparison, the choice is carried in different steps: we assume that consumers shortlist first products on the basis of few information on some relevant attributes, then they make the final choice, or they revise the shortlist. In particular, consumers first allocate the budget to a generic alternative; then shorlist alternatives; finally they choose the good in the basket of shortlisted alternatives, or revise the shortlist. We concentrate on the effect of limited information on the process of shortlisting in order to detect the combination effect of information on attributes on the second step of choice.

The limited information can result from business communication or advertising aimed at highlighting just few features of a product. We assume that on the basis of these few features, the decision maker puts a candidate good in his basket that, afterwards, will be subject to a closer comparison. As to firms, the problem is, then, to highlight the attributes that give more chances to draw the attention of the potential consumer, eventually increasing the chance to put firms' product in the basket of choice.

In order to define more clearly the object of our analysis, it is useful to formalize the shortlisting procedure.

Choice (in the deterministic case) is defined as a triple (Δ, C_δ, K), where Δ is the set of alternatives, C_δ are the consequences associated to each alternative and K is the choice criterion. We introduce now a *description* of alternatives. This can be defined as a vector whose elements are values attributable to each feature of an alternative in a space Ξ of attributes. A description $\xi_\delta \in \Xi$ is a vector of attribute values for $\delta \in \Delta$. K is, in this case, the criterion adopted to evaluate alternatives over their description. Usual criteria of choice in a multi-attribute setting are given by some sort of sum of described attributes compounded by a vector of weights, or by an evaluation of a distance between described alternatives and an ideal prototype. The consumer is then depicted as having a vector of weights ψ that permits to compound the described characteristics. In general, it is assumed that both alternatives and preferences are fully described: alternatives are defined by a complete description of all relevant characteristics and customers' preferences are defined by a complete weighting vector: in this sense, both alternatives and preferences are total objects, or completely defined elements. The problem of choice is then identified as: $((\Lambda, \Xi, \psi), C_\delta, K)$.

In our approach, consumers don't have a complete description of alternatives, but they have to rely on "rough descriptions" of objects. These *incomplete descriptions* can be formalized as follows. Let the vector of attributes be composed by n elements; an incomplete description is just a vector of n binary elements (the attributes) whose value is always 0, apart from the described k elements that take value 1. The simplified binary description of an alternative δ is a vector $\xi_\delta^\omega \in \Xi^{n,k}$ where n and k are respectively the total number of attributes and the number of them having a description, and $\Xi^{n,k}$ is the space of possible simplified descriptions of order k. Each alternative can be then described in $\begin{pmatrix} n \\ k \end{pmatrix}$ ways. For instance, as in one of the cases we are going to study, a car can be portrayed using two out of five attributes. ω represents exactly the attributes chosen to describe an alternative δ (a car can be described as being comfortable and safe). Obviously, from the point of view of their incomplete descriptions, more goods can be equivalent: this means simply that they enter in the shortlist.

To build their shortlist, decision makers use a simplified weighting system, by which they give equal weight to what they consider the most important attributes[1]. For instance, if to shortlist they consider (like in the examples we used for our empirical test) two attributes out of a total of five, their simplified weighting system will be a vector with three zeros and two $1/2$. Call ψ^{ω} this simplified weighting system, where ω is the subset of attributes that are chosen for shortlisting. The suggested criterion of choice will be to shortlist products whose inner product $\xi_{\delta}^{\omega}\psi^{\omega} = 1$. The intuition is that potential purchasers shortlist goods whose rough description matches the attributes considered most important.

We are now interested in the behaviour of a population Λ of potential consumers who react to different incomplete descriptions of a good. They will distribute their choices among the $\begin{pmatrix} n \\ k \end{pmatrix}$ possible descriptions of alternatives. Let $\lambda\xi_{\delta}^{\omega}$ be the number of individuals who choose the incomplete description ω for shortlisting (i.e. for which $\xi_{\delta}^{\omega}\psi^{\omega} = 1$); note that if the whole population make a choice, $\sum_{\omega} \lambda\xi_{\delta}^{\omega} = \Lambda$. We define the application $f^{n,k}$ on $\Xi^{n,k}$ as the function that gives $\lambda\xi_{\delta}^{\omega}$ for each element of $\Xi^{n,k}$.

We are interested in studying the empirical properties of such a function $f^{n,k}$, i.e. in analyzing how a population reacts to different incomplete descriptions of an alternative.

3 The survey

In order to analyze how a population reacts to descriptions of alternatives that make use of different combinations of attributes, we carried out a survey in the Italian provinces of Trento and Bolzano. In a first step we isolated, through open interviews on a sample of about 40 adults, the five most important attributes used when choosing 3 goods (a middle-capacity car, a mobile phone, a desktop computer) and 2 services (a dinner in a restaurant, and a 15 days holiday). We supposed that the task of choosing those alternatives was complex enough. To elicit the attributes people looked for in a choice context, we used a think aloud method. Then, we classified the attributes in order to harmonize answers. The attributes considered mostly important are presented in Table 1.

Table 1: List of attributes elicited for each alternative.

	Car	**Mobile phone**	**Restaurant**	**Desktop**	**Holiday**
A	comfort	handling	quality of service	hardware	resort
B	esthetics	functions	kitchen	performance	option package
C	safety	dimension	environment	software	accommodation
D	operating costs	esthetics	staff	size	entertainment
E	brand	brand	easy to reach it	assistance/warranty	easy to reach it

[1] It is easy to extend this formalization to a general case of uneven weights.

Elicited attributes were used to build up a second phase of the research, that consisted in a large survey on more than 1000 people, randomly selected, through a snowball sampling. The sample was equally divided between males and females (52%); other characteristics of respondents are summarized in Tables 2 and 3. The questionnaire was filled in public places, workplaces, and schools: this is the reason why elderly people are under represented. We asked people to fill a questionnaire divided in three parts. In the first one it was asked to mark for each alternative in a given range of prices the attribute in the list of five that had been considered the most important, would the respondent had to choose. In the second part people were asked to do the same thing with attributes coupled: there was then the possibility to mark one out of ten couples of attributes for each alternative ($AB, AC, AD, AE, BC, BD, BE, CD, CE, DE$). In the third part, some demo-social information was asked. In order to avoid the location effects, options were randomly ordered, both vertically and horizontally (the order of couples was inverted): we used indeed 10 versions of the questionnaire.

Table 2: The sample of respondents by age.

Age	% of respondent
less than 25	43.0
25–34	14.4
35–44	21.4
45–54	13.5
55–64	5.6
more than 65	2.2
	100

Table 3: Sample of respondent by profession and family in percentage

	nr members of the family			
Profession	1–2	3–4	≥ 5	Total
employee	24.6	13.2	3.6	41.5
self-employed	3.6	1.8	0.5	5.9
student	12.4	18.8	8.6	39.8
other	8.5	3.7	0.6	12.8
Total	49.2	37.5	13.3	100.0

Data were analyzed using both the usual descriptive statistics and the spectral analysis. Before presenting this second, methodologically more relevant part of our study, we highlight here some salient features of the choices, as they emerge form descriptive statistics.

- **car**: women and the elderly people care safety. Youngest people give importance to aesthetics and they care operating costs more then middle aged people. When attributes are coupled, the preferred couple is given by safety and operating costs; in general women prefer couples that contain safety while men those containing operating costs;
- **mobile phone**: with the increasing of age, we observed a shift in preferred attributes from functions to handling. In the choice of couples, coupling handling and functions

appears particularly attractive while esthetics grows in importance and attracts mostly preferences, when coupled with functions or handling.

- **restaurant**: with the increasing of age, up to a given point, preferences shift from the environment to the kitchen; then, above a certain age, it emerges that service and easy to reach become more important. In the choice of couples, we note strong preferences when kitchen is paired with the personnel assistance and when kitchen is coupled with the service;

- **desktop**: performances are mostly important for young people, students and big families, while - probably due to better technical skills - assistance less important. On the contrary, assistance becomes more important among self employed and elderly people. In the choice of couples, performances and components pair well and are important among young people and students while the couple which pairs performance and assistance is important for the other classes;

- **holiday**: among elderly people, the resort decreases in importance in favour of accommodation and the easy to reach, while resort and entertainment are important among people with a higher education. In the choice of couples, the most important is, by far, the one that joins locality with accommodation. The couple "resort and entertainment" is preferred by people with higher educational qualification.

The descriptive analysis of our survey helps highlight particular features of the choice process in different populations. The spectral analysis, instead, would help us separate the weight given to single attributes, from the effect of their coupling. By describing the working method we will focus on a particularly interesting case: the evaluation of a car.

4 Noncommutative harmonic analysis

In this paper, we make use of noncommutative harmonic analysis to examine data on preferences gathered through the survey described in the previous section. Being a generalization of classical spectral analysis this mathematical framework is also known as "generalized spectral analysis" or else as "discrete Fourier analysis", and it has been widely used in time–series analysis (see Chatfield (1975)) and in computer, engineering and natural sciences. It is a non–model based approach to data analysis and was formulated in a general group theoretic setting by Diaconis (see Diaconis (1988) and Diaconis (1989)), who extended the classical spectral analysis of time series to the analysis of discrete data having a noncommutative structure.

The main idea of spectral analysis is that often data have natural symmetries, which are hidden in the existence of a symmetric group (which is obviously non commutative, wherefore the name of noncommutative harmonic analysis) for the domain of the data. The leading principle of spectral analysis is the interpretation of data through its decomposition according to these symmetries. New efforts have been made in order to apply spectral analysis to non–time series in the political sciences, above all in the analysis of voting.

Recently, Lawson and Orrison (2002) and also Lawson et al. (2003) have introduced a generalization of spectral analysis as a new instrument for political scientists; they used the powerful machinery of spectral analysis to analyse political voting data. In particular, they analysed votes of the nine judges of the United States Supreme Court (Warren Court 1958–1962, Burger Court 1967–1981, Renquist Court 1994–1998) and detected influential coalitions.

The idea followed by Lawson and Orrison (2002) is to consider political voting data as elements of a mathematical framework; then the features of that framework can be used to work out natural interpretations of the data. The mathematical framework corresponding to voting data has many components, each of which encapsulates information on particular "*coalition effects*"; the decomposition of data with respect to these components provides the identification of **influential coalitions**.

Rensi and Zaninotto (2004) use this machinery to analyse the *power of combination* in simplified preferences.

In order to apply noncommutative harmonic analysis to our setting, let $X = \{x_1, ..., x_n\}$ be a finite set of n elements and $f : X \to \mathbb{C}$ a complex valued function on X. In our contest the set X represent the set of choice options $\Xi^{n,k}$ both single and in couple, while f is the frequency vector of the generic alternative, $f^{n,k}$, as defined in Section 2. Let P the vector space of the complex valued function under examination. f is an element of P. We can consider P as an *ambient space*. We may decompose P into a direct sum of invariant subspaces encapsulating important properties of the data and projecy f in these subspaces.

P may always be decomposed into a direct sum

$$P = P_1 \oplus \ldots \oplus P_h$$

for some positive integer h. In particular, each function $f \in P$ may be written uniquely as a sum

$$f = f_0 + \ldots + f_h$$

with $f_i \in P_i$.

We define the set of five attributes for each good $\Xi^{5,1} = \{A, B, C, D, E\}$ and the ten couples of attributes for each good $\Xi^{5,2} = \{AB, AC, AD, AE, BC, BD, BE, CD, CE, DE\}$. We define, as in Section 2, the vector of choices of a population over a single attribute description as the function $f^{5,1}$, while $f^{5,2}$ is the function for double attributes description.

Vector spaces for functions $f^{5,1}$ and $f^{5,2}$ may be decomposed in

$$P^{5,1} = P_0^{5,1} \oplus P_1^{5,1}$$

$$P^{5,2} = P_0^{5,2} \oplus P_1^{5,2} \oplus P_2^{5,2}$$

We may project the function $f^{5,1}$ and $f^{5,2}$ onto these invariant subspaces and obtain

$$f^{5,1} = f_0^{5,1} + f_1^{5,1}$$

$$f^{5,2} = f_0^{5,2} + f_1^{5,2} + f_2^{5,2}$$

The projections of functions onto the subspaces capture, in the single choice, the mean effect $f_0^{5,1}$ and the first order effect $f_1^{5,1}$ while in the choice of couples, they capture the mean effect $f_0^{5,2}$, the first order effect $f_1^{5,2}$ and the second order effect $f_2^{5,2}$. To interpret these effects we adapted to our purpose Lawson and Orrison (2002) and Rensi and Zaninotto (2004).

In the single choice, the first order effect $f_1^{5,1}$ tells us if an attribute is chosen on average. If the first order effect is positive it means that the attribute is chosen on average. In this simple case, noncommutative harmonic analysis doesn't add any information.

More important is the analysis of first and second order effects of $f^{5,2}$. In this case, $f_1^{5,2}$ (the first order effect) is a sort of average effect: it tells us if an attribute is chosen along the pairs to which it is coupled. This effect can be more clearly understood if, using *Mallow's Method* (Mallows, 1957), we sum every single first order effect of an attribute: in this way, we obtain the contribution of a single attribute in a generic couple. To do this, we calculate the inner product between the function $f_1^{5,2}$ and a function $f_H^{5,2} \in P^{5,2}$. This function $f_H^{5,2}$ identifies the elements of $f_1^{5,2}$ "containing" H with 1 and those "non containing"H with 0, e.i. if H=A (we want to consider the attribute A) the $f_H^{5,2}$ will be

$$f_A^{5,2} = (1, 1, 1, 1, 0, 0, 0, 0, 0, 0).$$

The contribution of the attribute A is $f_{1,A}^{5,2} = f_1^{5,2} f_A^{5,2}$. [2]

The second order effect $f_2^{5,2}$ can tell us something about the *power of combination* between the two attributes. This power may be interpreted as a sort of *complementarity* between attributes. If two attributes have a positive second order effect, it means that pairing two attributes in a couple would result in a higher choice than what had been observed just taking into account the power of attraction of single attributes. We can say that in this case attributes show a reinforcement effect when paired in a couple. The reverse is true if the second order effect is negative: in this case, by pairing attributes their power of attraction is weakened.

In order to help the interpretation of first and second order effects, they can be simultaneously visualized in the following cross table (see Table 4).

Table 4: First and second order effects

		$f_1^{5,2}$	
		−	+
$f_2^{5,2}$	+	The couple is not chosen on average but has a reinforcement effect. It is better to preserve the couple in order to attract more preferences.	The couple is chosen on average and has a reinforcement effect. Couples in this box are winning.
	−	The couple is not chosen on average and has no reinforcement effect. It is better to avoid using those attributes in communication	This couple is chosen on average, while it does not have any reinforcement effect. Unpair the attributes, keeping the strongest ones isolated.

5 An application: the choice of a car

In order to understand how the method can be useful in analyzing data, giving light to some "hidden" pattern of phenomena under examination, let us start by showing the frequency tables on the observed preferences for car attributes. The observation of descriptive statistics suggested

[2]We cannot directly compare the first order effect in the single attribute choice with the first order effect for the same attribute in a couple, as it results form the application of Mallow's Method, because in this case the attribute is counted independently from the order in which it appears while in the single attribute choice we have only asked what had been the *most important* item used for choosing.

us to compare choices by gender: males and females attitudes towards choice appear to be sensibly different, suggesting the existence of specific cognitive evaluation for single attributes and their joint presentation.

We begin with the table of single choice preferences in order to get $f^{5,1}$.

Table 5: Preferences on single attributes in the car evaluation

		$f^{5,1}$	
	Single Attribute	Male	Female
A	Comfort	66	57
B	Esthetics	83	86
C	Safety	116	191
D	Operating costs	148	161
E	Brand	76	43
	Total	489	538

We may project the two functions $f^{5,1}$ onto the invariant subspaces decomposition $P^{5,1} = P_0^{5,1} \oplus P_1^{5,1}$, getting:

Male

$$\begin{pmatrix} 66 \\ 83 \\ 116 \\ 148 \\ 76 \end{pmatrix} = \begin{pmatrix} 97.8 \\ 97.8 \\ 97.8 \\ 97.8 \\ 97.8 \end{pmatrix} + \begin{pmatrix} -31.8 \\ -14.8 \\ 18.2 \\ 50,2 \\ -21,8 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \\ E \end{matrix}$$

$$f^{5,1} = f_0^{5,1} + f_1^{5,1}$$

Female

$$\begin{pmatrix} 57 \\ 86 \\ 191 \\ 161 \\ 43 \end{pmatrix} = \begin{pmatrix} 107.6 \\ 107.6 \\ 107.6 \\ 107.6 \\ 107.6 \end{pmatrix} + \begin{pmatrix} -50.6 \\ -21.6 \\ 83.4 \\ 53.4 \\ -61.6 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \\ E \end{matrix}$$

$$f^{5,1} = f_0^{5,1} + f_1^{5,1}$$

Analysing $f_1^{5,1}$ we note that operating costs D are equally important for both genders. Females give importance to safety more than males. C is more important than D for females. Males preferences are more distributed along attributes than females; in particular, females preferences are concentrated on the two more important attributes.

As to choice of couples, let us start by showing the table of frequencies:

Table 6: Frequency of preferences on couples of attributes in car evaluation

		$f^{5,1}$	
	Couple of Attributes	Male	Female
AB	comfort and esthetics	38	41
AC	comfort and safety	69	89
AD	comfort and operating cost	54	44
AE	comfort and brand	23	15
BC	esthetics and safety	45	72
BD	esthetics and operating cost	64	53
BE	esthetics and brand	39	14
CD	safety and operating cost	83	155
CE	safety and brand	38	35
DE	operating cost and brand	38	18
	Total	491	536

We may project the two functions $f^{5,2}$ onto the invariant subspaces decomposition $P^{5,2} = P_0^{5,2} \oplus P_1^{5,2} \oplus P_2^{5,2}$.

As to males we obtain:

$$f^{5,2} = \begin{pmatrix} 38 \\ 69 \\ 54 \\ 23 \\ 45 \\ 64 \\ 39 \\ 83 \\ 38 \\ 38 \end{pmatrix} = \begin{pmatrix} 49.1 \\ 49.1 \\ 49.1 \\ 49.1 \\ 49.1 \\ 49.1 \\ 49.1 \\ 49.1 \\ 49.1 \\ 49.1 \end{pmatrix} + \begin{pmatrix} -7.6 \\ 8.7 \\ 10.1 \\ -23.6 \\ 9.4 \\ 10.7 \\ -22.9 \\ 27.1 \\ -6.6 \\ -5.3 \end{pmatrix} + \begin{pmatrix} -3.5 \\ 11.2 \\ -5.2 \\ -2.5 \\ -13.5 \\ 4.2 \\ 12.8 \\ 6.8 \\ -4.5 \\ -5.8 \end{pmatrix} \begin{matrix} AB \\ AC \\ AD \\ AE \\ BC \\ BD \\ BE \\ CD \\ CE \\ DE \end{matrix}$$

$$f^{5,2} = f_0^{5,2} + f_1^{5,2} + f_2^{5,2}$$

and as to females:

$$f^{5,2} = \begin{pmatrix} 41 \\ 89 \\ 44 \\ 15 \\ 72 \\ 53 \\ 14 \\ 155 \\ 35 \\ 18 \end{pmatrix} = \begin{pmatrix} 53.6 \\ 53.6 \\ 53.6 \\ 53.6 \\ 53.6 \\ 53.6 \\ 53.6 \\ 53.6 \\ 53.6 \\ 53.6 \end{pmatrix} + \begin{pmatrix} -19.9 \\ 37.1 \\ 10.1 \\ -52.6 \\ 34.1 \\ 7.1 \\ -55.6 \\ 64.1 \\ 1.4 \\ -25.6 \end{pmatrix} + \begin{pmatrix} 7.3 \\ -1.7 \\ -19.7 \\ 14.0 \\ -15.7 \\ -7.7 \\ 16.0 \\ 37.3 \\ -20.0 \\ -10.0 \end{pmatrix} \begin{matrix} AB \\ AC \\ AD \\ AE \\ BC \\ BD \\ BE \\ CD \\ CE \\ DE \end{matrix}$$

$$f^{5,2} = f_0^{5,2} + f_1^{5,2} + f_2^{5,2}$$

In order to gain a better understanding of data, we can use the following graph (Figure 1), that plots the frequencies of choices by first and second order effects, as described in Section 4.

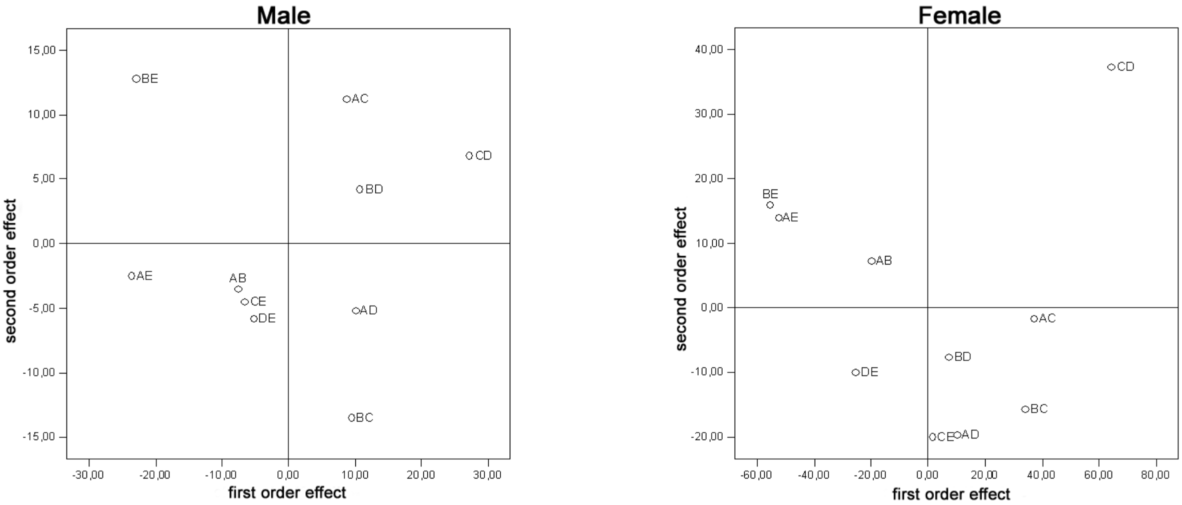

Figure 1: The first and second order effects

As to males, both first and second order effects are positive in three cases, while for females this happens only once. Females concentration of preferences was already evident in single attribute choices: they want to find the right trade-off between safety (C) and operating costs (D). As a consequence, in the choice of pairs, the couple CD is chosen on average and reveals, at the same time, a strong reinforcement effect. The couple AC (comfort and safety) and the couple BD (esthetics and operating cost) are equally chosen on average but they show a reinforcement effect only for men. Males appear to pair equally well comfort and safety, and costs and aesthetics. Comfort and operating cost (C and D) and esthetics and safety (B and C) match poorly: we find them in the 2^{nd} box: their pairing has a weakening effect for both genders. It could be that the two attributes transmit contrasting feelings about the product: while positively evaluating comfort or low operating costs singularly considered, people receive discordant cognitive signals from their pairing. In this case, it would be better to separate communications on attributes in order to avoid to weaken the strongest one. The couple CE (safety and brand) is chosen on average only by females. The couple DE (operating cost and brand) is posited in the 3^{rd} box: it is poorly chosen on average and shows a weakening effect for both genders. The couples AB and AE (comfort and esthetics, comfort and brand) are not chosen on average by both genders, but females choices show a reinforcement effect. The last couple BE (esthetics and brand) shows a reinforcement effect in both genders: this suggests to preserve the couple in order to concentrate the preference of a minority.

It should be clear from our description of first and second order effects that sensible differences exist between males and females behaviour. We can deepen the first order effect by studying the contribution of each single attribute in a generic couple. We used the *Mallow's Method* to evaluate this contribution. Results are presented in Table 7.

Attributes C and D have a positive contribution in a generic couple. The effect of C is greater among females, while the effect of D is analogous for both genders. The other three attributes show negative contributions. Absolute values are greater among females because strong preference for attributes prevails. We note that there are not wide differences between the first order effect in single attribute choice and the contribution of a single attribute in the choice of couples using Mallow's method.

Table 7: The first order effect using the Mallow's method

	Attribute	$f_{1,h}^{5,1}$ Male	Female
A	comfort	-12.4	-25.4
B	esthetics	-10.4	-34.4
C	safety	38.6	136.6
D	operating cost	42.6	55.6
E	brand	-58.4	-132.4

Focusing on the second order effect, it is possible to highlight how males tend not only to use more attributes, but also to link the attributes evaluation. Among males, several attributes are coupled with positive effects while among females a more straight model of choice prevails. Here, subgroups of attributes (A,E,B, and C,D) show positive second order effects, but are reciprocally isolated.

This different pattern of behaviour is clear in the graphical representation presented in figure 2, that depicts the positive link among attributes, as they result from a positive second order effect.

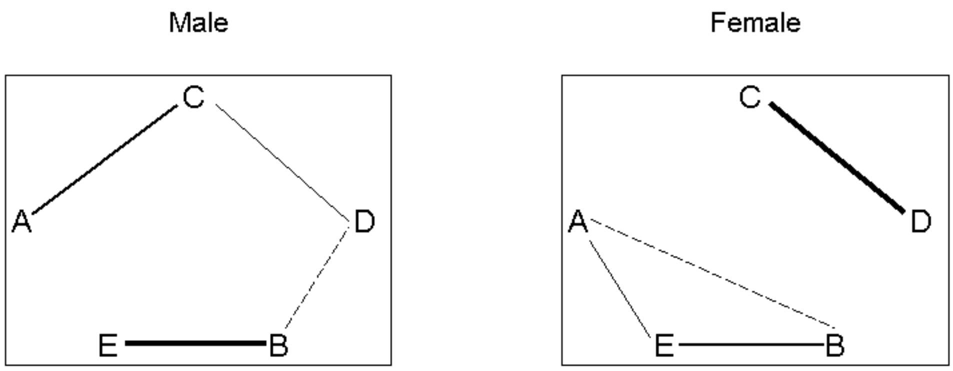

Figure 2: The second order effect

The picture is obtained by linking the attributes with a line that have a reciprocal positive second order effect. The thickness of the line, represents the intensity level of the bond (i.e. the relative height of second order effect); the absence of a link means that the second order effect between a given pair is negative. The figure representing females behaviour shows that there are two groups of reciprocally reinforcing attributes: the most and the less preferred ones. The male case is more interesting: the bond between stronger attributes is not the most important one. The most important bond is instead between two weak attributes: esthetics and brand (E and B). Strong attributes show a positive reinforcement effect if they are reciprocally coupled, or if they are coupled with weaker attributes (except for the brand E).

6 Conclusions

The case we presented in the previous section was aimed to give an example of the use that it is possible to make of noncommutative harmonic analysis. In particular, we used it to extract information from data collected in our survey on preferences, assuming that consumers shortlist choices on the basis of an incomplete description of attributes. A comparative analysis of first and second order effects through different goods, or among different consumers of the same goods, can help understand factors that motivate the choice, and detect if and when attribute descriptions complement or not. This analysis was presented in detail in the case of a car choice.

¿From a practical point of view, a development of such a kind of analysis could help understand which attributes need to be highlighted in advertising or communication, in order to increase the probability of being shortlisted by potential consumers. The first order effect helped understand which attributes are chosen on average through the attributes to which they are coupled. The analysis of the second order effect allowed us to understand which attributes, if coupled, show reciprocal reinforcement effect.

Obviously, much should be done to increase the prediction power of the method. An obvious step onward should consist in finding regularities in first and second order effect.

¿From a psychological point of view, it could be interesting to understand why some attributes reciprocally reinforce or weaken. In particular, it could be interesting to understand whether second order effects between attributes were due to an information problem (people tend to skip attributes that don't add information on a given good, and vice versa), or they are due to the cognitive process that people adopt to evaluate the utility of an alternative. From this point of view, the method we have presented here can offer a row material for a deeper analysis.

References

Chatfield, C. (1975). *The analysis of time series. an introduction*. Chapman and Hall, London.
Diaconis, P. (1988). *Group representations in probability and statistics*. IMS, Hayward, CA.
Diaconis, P. (1989). A generalization of spectral analysis with application to ranked data. *The Annals of Statistics*, *17*(3), 949–979. (The 1987 Wald Memorial Lectures)
Diaconis, P., & Rockmore, D. N. (1990). Efficient computation of the fourier transform on finite groups. *J. Amer. Math. Soc.*, *3*(2), 297–332.
Lawson, B. L., & Orrison, M. E. (2002). *Analyzing voting from a new perspective: Applying spectral analysis to the u.s. supreme court.* Presented at the annual meeting of the American Political Science Association, Boston, August 29.
Lawson, B. L., Orrison, M. E., & Uminsky, D. T. (2003). *Discrete analysis of voting in committees.* Presented at the annual meeting of the Midwest Political Science Association, Chicago, April 3–6.
Mallows, C. L. (1957). Non-null rankings models i. *Biometrika*, *44*, 114–130.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psycology: Learning, Memory and Cognition*, *14*(3), 534–552.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press, Cambridge.
Rensi, S., & Zaninotto, E. (2004). Simplified preferences, voting, and the power of combination. *DISA WP, University of Trento*, *90*.