CAHIER DU LAMSADE

Laboratoire d’Analyse et Modélisation de Systèmes pour l’Aide à la Décision
(Université Paris-Dauphine)
Unité de Recherche Associée au CNRS n° 825

COMPENSATORINESS OF PREFERENCES
IN MATCHING AND CHOICE

CAHIER N° 122
janvier 1994

Vincent MOUSSEAU ¹


¹ LAMSADE, Université Paris-Dauphine, Place du Maréchal De Lattre de Tassigny, F-75775 Paris Cedex 16.
(NON-)COMPENSATION DES PRÉFÉRENCES
ÉVIDÉES PAR GÉOMÉTRIE NETTE
CONTENTS

Résumé ......................................................................................... i
More recently, [Tversky et al. 88] showed that the same kind of violation may occur within the context of riskless decisions. In this paper, two different questioning modes were studied in a bi-criterion context: direct binary choice between alternatives and matching (matching questions consist in proposing two alternatives to the DM, one of the evaluations of an alternative being unfixed; the DM is to determine this evaluation in order to be indifferent between the two alternatives). The authors showed a prominence effect which states that "the more prominent attribute looms larger in choice than in matching", in other words that "preferences induced from choice are likely to be closer to the lexicographic ordering than those induced by matching".

Our work takes [Tversky et al. 88] as a starting point and aims at confirming and extending the obtained results. Unlike the above mentioned study that tests the prominence effect on a small number of specific pairs of alternatives, we want to analyse this effect on the overall preference structures of DMs. The knowledge of these overall PSs allows us to evaluate the "amount of compensation" the DM accepts when answering to choice and matching questions.

Our analysis of the phenomenon initially observed by [Tversky et al. 88] differ from the one proposed by these authors. We use the general tendency of the empirical observations in order to induce a rule concerning the questioning mode to be used in importance parameters elicitation techniques. The questioning mode of an elicitation techniques is to be chosen in regard to the aggregation procedure used to model preferences. The information obtained from the DM should to be consistent with the use of this information in the aggregation procedure. Consequently, if the chosen aggregation procedure is of a non-
2. (Non-)Compensatoriness of preference structures

The notion of (non-)compensatoriness of preference structures is intuitively linked to the possibility to resort to substitution rates in the construction of preferences that take all criteria into account. These substitution rates allow to compensate a disadvantage on a particular criterion by a sufficient advantage on another criterion. When an aggregation procedure uses such ideas to solve conflicts between criteria, the induced preference structure is said to be compensatory, otherwise it is said to be non-compensatory. Similarly, preferences expressed by a DM denote a preference structure of more or less compensatory nature.

2.1. Definitions

A first formal definition of this notion has been proposed by [Fishburn 76]: Let us define by
\[ P \text{ and } I \text{ the preference and indifference relations on a product set } X \times X \times X \times X \]
More recently [Roy & Mousseau 92] proposed a definition allowing the analysis of compensation on each criterion individually (this definition is grounded on a new formalism of the notion of relative importance of criteria). However the (non-)compensatoriness of a preference structure is a notion that deserves to be studied in further details. Firstly no large consensus emerges concerning its definition. Moreover the available definitions only define extreme cases and do not allow to precise to which level a preference structure is (non-)compensatory.

2.2. Construction of a (non-)compensatoriness index

Total non-compensation is an extreme situation never reached in real world decision contexts. Let us define an index aiming at "measuring" the (non-)compensatory aspects of a preference structure $\Psi = (X,P,\Pi_p)$. $\Psi$ is even more non-compensatory since two pairs of alternatives having the same preferential profile on criteria are linked with the same
In this definition, \( \lambda \) is the number of possible preferential situations on overall level (when preferences are modeled using a (1,P) structure, three preferential situations are possible between two alternatives: aPb, alb and bPa). This index is built such that \( \text{nc}(\Psi) = 1 \) means that \( \Psi \) is totally non-compensatory (lexicographic order); a decreasing value for \( \text{nc}(\Psi) \) means greater possibilities of compensation in \( \Psi \). Thus the index \( \text{nc} \) allows us to compare the (non-)compensatoriness of two preference structures; this index has only an ordinal signification (no cardinal use of this index will be made).

3. Empirical scheme

Our aim in this experiment is to highlight the violation of procedural invariance and to exhibit a link between the use of a specific questioning mode (choice and matching in our case) and the (non-)compensatoriness of the elicited preference structures. We will deal with a bi-criteria context. The scale on both criteria will be discretized on a four level scale \( \{A,B,C,D\} \) where \( A \geq P_i \geq B \geq C \geq D \) \( i=1,2 \).

3.1. The two questioning modes

The experiment consists in the exhaustive evaluation of the preference structure using two different questioning modes:

**Choice**: this type of questions consists in a hollistic comparison of two alternatives defined by their evaluations on all criteria. Subjects are to choose between an indifference situation or a preference in favor of one of the alternatives. As we deal with a bi-criteria context, let us denote by \((x_1,x_2)?(y_1,y_2)\) such a question.

**Matching**: this type of question consists in proposing two alternatives \((x_1,x_2)\) and \((y_1,y_2)\) leaving \(y_2\) unfixed. Subjects are to determine \(y_2\) in order to obtain indifference between the two alternatives. Let us denote by \((x_1,x_2)?(y_1,?)\) such a question. As in our experiment answers are given on a continuous scale, it is necessary to analyse the answers in terms of the discretized four level scale \((A \rightarrow D)\). In order to do so we use an indifference threshold \( q \) which represents the minimum discernable difference between two evaluations (see [Roy & Vincke 87]). This threshold is elicited beforehand and used to compare the value \(y_2\) to the four levels of the discretized scale. The nine possible situations are represented in figure 1.
3.3. Experimental framework

The real world context concerns the evaluation of firms by young executives working in computer science companies so as to postulate in these firms. The proposed firms differ in the annual salary and the job interest. The scales of both criteria (salary and job interest) are discrete and made of four levels of evaluation.

Criterion 1: The job interest is evaluated on a qualitative scale that is built with subjects. Each level is defined by linguistic terms. Instructions were given to subjects: they had to build this scale so as to perceive the "distance" between consecutive levels as equivalent. It was then checked that all levels are separated by a strict preference.

For example, the scale of one of the subject was:

A: Very interesting job, formative, no repetitive aspect, evolutive and large independence at work.
B: Interesting job, slightly repetitive but formative, rather evolutive, good independence at work.
Postulate 2 (monotonicity): \[ \forall x, y, z, t \in X \begin{cases} xPy \text{ and } tA_p x \Rightarrow tPy \\ xPy \text{ and } yA_p z \Rightarrow xPz \end{cases} \]

The interpretation of this postulate is the following: "when an assertion aPb is established, increasing the evaluation of \( x \) or decreasing the evaluations of \( y \) leaves the overall preference relation between \( x \) and \( y \) unchanged.

Postulate 3 (a partial preference is valid on the overall level):
\[
\begin{align*}
& xIy \text{ and } yA_p z \Rightarrow xPz \\
& xIy \text{ and } tA_p x \Rightarrow tPx \\
\end{align*}
\]
i.e.
\[
\begin{cases} I \Delta_p \subset P \\
\Delta_p I \subset P \end{cases}
\]

As in the preceding postulate, this postulate enables us to induce assertion from previously determined assertions. In concrete terms, if \((x_1, x_2)I(y_1, y_2)\) then we have:
\[
\begin{align*}
& \forall (z_1, z_2) \in X \ (z_1, z_2)A_p (x_1, x_2) \Rightarrow (z_1, z_2)P(y_1, y_2) \\
& \forall (t_1, t_2) \in X \ (y_1, y_2)A_p (t_1, t_2) \Rightarrow (x_1, x_2)P(t_1, t_2) \\
\end{align*}
\]

3.5. Data acquisition software

According to the preceding postulates, we build an algorithm aiming at reducing the number of questions to be posed to the DM in order to elicit his overall preference structure. Let us ground our algorithm on the oriented graph \( G=(C,U) \) where \( C \) is composed of every pair \((x,y)\) with \( x \neq y \) whose comparison is not determined by postulate 1 and \( U \) represents
Figure 2

Your choice:

I prefer the firm Y to the firm X.
Choosing between firm X and Y is indifferent.
answering time could be recorded. The average answering time on the whole subject group is 24 seconds for matching questions and 30 seconds for binary comparisons. A T Student test shows that these two mean values differ significantly with α≤0.01. This difference may be explained by the fact that only 3 answers are possible in choice questions while subjects are to answer to matching questions on a continuous scale. Moreover we observed during the interviews that subjects seem to test a few binary comparisons in order to answer to matching questions.

4.1. Prominence effect

Let us recall that the prominence effect states that preferences elicited by binary comparisons are closer to a hierarchically ordered set than those elicited by matching questions.
In this experiment, the mean value for $p(\Psi_c)$ (respectively $p(\Psi_m)$) is 0.800 (resp. 0.663) and its standard deviation is 0.116 (respectively 0.133). A Wilcoxon test leads to accept that the mean values of $p(\Psi_c)$ and $p(\Psi_m)$ are significantly different with $\alpha \leq 0.01$. A T Student test leads to the same conclusion with $\alpha \leq 0.01$. We can thus conclude that our experiment confirms the prominence effect.

4.2. Contingent compensation effect

The contingent compensation effect is closely related to the prominence effect and states that elicited preferences reveal more possibilities of compensation when the questioning mode is matching rather than binary comparison. So as to measure the possibilities of compensation inherent in a DM’s preference structure, we use the index nc built in section 3.2. Let us recall that this index varies between $\frac{1}{3}$ and 1 and the higher the value for nc($\Psi$), the less possibilities of compensation exist in $\Psi$ (nc($\Psi$)=1 corresponding to a lexicographic order).

We compute the values of this index for both questioning modes and all subjects. Results are synthetized in figure 5 in which each point represents a subject. The scatterplot is located below the bisectrix, i.e. the values for nc($\Psi_c$) are higher than those of nc($\Psi_m$) for a large majority of subjects.
In our sample, the mean value across subjects for the computed value of the index nc is 0.535 (respectively 0.685) when preferences are elicited by matching questions (respectively choice questions) and the standard deviation is 0.145 (respectively 0.144). A Wilcoxon test leads to conclude to a significant difference between the mean values for nc(Ψ_o) and nc(Ψ_m) with α≤0.01. A T Student test leads to the same conclusion with α≤0.01. We can thus conclude that our experiment confirms the contingent compensation effect.

5. Discussion

The experimental results show a strong failure of the procedure invariance principle: we observed, in our sample, a significant divergence among preferences elicited through two different questioning modes (choice and matching questions). More precisely, subjects answered in such a way that matching questions elicit preference structures in which more possibilities of compensation are allowed and in which the preponderant criterion looms larger
preferences are supposed to remain constant but to be distorted during the elicitation process. Elicitation techniques orient preferences in certain directions, introduce noises ... However it is assumed that, when the analyst and the DM are sufficiently careful, have enough time and use different elicitation methods, obtained preferences should converge on true preferences.

Lability of preferences may also be analysed following a constructivist approach. In this case, preferences are not assumed to be pre-existent to the modelling process. [Bell et al. 88] emphasize that "it is a planotic myth that latently probabilities and utilities really exist deep down and that the analyst merely has to cut away the fat in order to display the pre-existing structure". Observed preferences are considered as a construct of the elicitation process and analysed as a result of interactions between the DM and the elicitation tool. This does not mean that DMs are free from any opinion but the constructs refer to basic attitudes, values and opinions that cannot be observed directly and to which we have access only through the filter of an elicitation procedure. Lability of preferences results from differences in the interactions between DMs and elicitation tools.

In a descriptivist interpretation, the violation of the procedure invariance principle is then analysed as biases in the elicitation procedures (see [Hershey et al. 82]). In our case, choice questions are supposed to be biased in such a way that DMs tend to "overweight the more important criterion" while matching questions "push" DMs in the opposite direction. The "true pre-existing preferences" are supposed to be in between and the use of different questioning modes should lead to converge to these true preferences.

In the constructivist approach, DMs' preferences are not viewed as totally pre-formed; they are built (at least partially) during the modelling process. In this framework, elicited preferences are the result of an interaction between the DM and the elicitation tools; in this sense, this tool is not to be considered as neutral. Following this approach, the experimental results described in this paper may be used in two distinct ways according to the objective to be pursued. [Tversky et al. 88] obtained similar results and proposed a model (the contingent tradeoff model) that accounts for the observed divergences among the preferences elicited by different questioning modes in which "the tradeoffs among inputs depend on the nature of the output". In this case, the pursued objective is to account for the observed phenomenon through an explicative model.

We propose an alternative use of the same observed phenomenon. Our goal is to use the general tendency of the empirical observations in order to induce a rule concerning the questioning mode to be used in importance parameters elicitation techniques. We observe that binary comparisons induce a preference structure of a more non-compensatory nature than the one elicited with matching questions for the same DM. It seems to us that the choice of a questioning mode for elicitation techniques should be made in regard to the aggregation
procedure used to model preferences. It is crucial for the information obtained from the DM to be consistent with the use of this information in the aggregation procedure. Consequently if the chosen aggregation procedure is of a non-compensatory nature (lexicography, majority rule), then the use of binary comparisons seems to be more adequate than matching questions: choice questions will induce an information that is likely to be more consistent with a non-compensatory aggregation rule. Conversely, matching questions will be more adapted to elicit preferences when a compensatory aggregation is used.

6. Conclusion

The experimental study reported in this paper shows that DMs strongly violate the procedure invariance principle. Our results parallel and enhance those of Tversky et al.
References


Mousseau V. (1992) 'Are judgments about relative importance of criteria dependent or independent on the set of alternatives? an experimental approach', Université Paris-Dauphine, Cahier du LAMSADEN°111.


