Cahier d’Études et de Recherche / Research Report

The effect of multicriteria conflict on matching-elicited preferences

Stéphane Deparis, Vincent Mousseau, Meltem Öztürk and Caroline Huron

CER13–02
Novembre 2013
The effect of multicriteria conflict on matching-elicited preferences.

Stéphane Deparis

*IBM Research Ireland, Technology campus, Mulhuddart, 15, Ireland*

Vincent Mousseau∗

*Laboratoire Génie Industriel, Ecole Centrale Paris, Châtenay-Malabry, 92290, France*

Meltem Öztürk

*LAMSade, Université Paris Dauphine, Paris 75016, France*

Caroline Huron

*INSERM Cognitive Neuroimaging Unit, CEA Neurospin, Gif-sur-Yvette, F91191, France*

Abstract

In this research, we focus on multicriteria preference elicitation by matching. In this widely employed task, the decision maker is presented with two multicriteria options, a and b, and must assess the value on one criterion for b, left blank, so that she is indifferent between the two options. A reverse matching, normatively equivalent, can be created by integrating the answer to the description of b and letting the DM adjust a performance value on the previously totally specified option a. Theory assumes that isopreferences resulting from the forward and reverse matchings are identical, but they empirically differ in a systematic direction. We investigate the effect of the multicriteria conflict, or trade-off size, on the asymmetry of judgement between forward and backward matchings. We observed that the asymmetry of judgement is

∗Corresponding author

*Email addresses: stephane.deparis@ie.ibm.com (Stéphane Deparis), vincent.mousseau@ecp.fr (Vincent Mousseau), ozturk@lamsade.dauphine.fr (Meltem Öztürk), caroline.huron@orange.fr (Caroline Huron)*

Preprint submitted to Elsevier November 11, 2013
increased both by multicriteria conflict, and by asking deteriorating rather than improving judgements at both steps of the double matching. We derive some implications for the practice of preference elicitation.

Keywords: multicriteria decision, behavioural decision analysis, preference elicitation, multicriteria conflict, matching, experiment

1. Introduction

Supporting decision makers (DMs) involved in decision processes often requires making use of preference models which incorporate the value systems and judgements of these DMs. Hence the concept of preference is crucial in decision making, and accounts for the way DMs evaluate and compare alternatives. The field of preference modelling has steadily grown over the last fifty years and numerous models have been proposed, such as utility-based models (Keeney and Raiffa (1976)) or outranking-based models (Roy (1996)). Preference models define rationality from a normative point of view, that is, norms that the decision maker (DM) should conform to.

Implementing preference models to support DMs involved in decision processes requires incorporating their judgements in the model. This requires to acquire preferences through an interaction with DMs, and to integrate these preferences in the model: this is called preference elicitation. Numerous preference elicitation methodologies have been proposed in the literature (Jacquet-Lagreze and Siskos (1982), Mousseau and Slowinski (1998), Bana e Costa and Vansnick (1994), Edwards and Barron (1994)) so as to capture DMs points of view and integrate them into a specific preference model. It is clear that the preference elicitation process is crucial so as the preference model to faithfully represent DMs opinions, and lead to recommendations which can be viewed as reasonable and helpful by the DMs.

However, preference elicitation methodologies implicitly hypothesize rationality principles to which DMs do not always conform. This can lead to elicited preference model which do not faithfully represent DMs judgments. It is therefore important to fully understand decision behavior through empirical studies in order to avoid misinterpretation of DMs preference statements
and to avoid decision biases. For instance, preference elicitation methods usually assume “procedure invariance which requires strategically equivalent methods of elicitation to yield the same preference order”. (Tversky et al. (1990)).

Violations of this procedure invariance principle have been observed when using matching vs. pairwise comparison tasks (e.g., Fischer and Hawkins (1993), Fischer et al. (1999), Tversky et al. (1988), Willemsen and Keren (2002)). In pairwise comparisons, the DM compares two alternatives and says if she prefers one to the other or if she is indifferent between them. In matching tasks, two alternatives are presented to the DM, one with assigned evaluations on all criteria and the other missing a value for one criterion. She is then asked to provide the missing value which makes the two alternatives indifferent. The literature provides explanations of such preference reversals which may involve the prominence effect (e.g., Slovic (1975), Slovic (1995)), the choice of the adjusted criterion (Slovic et al. (1990), Fischer and Hawkins (1993)) or the framing effect (Tversky et al. (1988)).

A positive feature of matching compared to other procedures is that it provides rich information on trade-offs from a limited number of questions (see Carmon and Simonson (1998), for a discussion on this topic). For compensatory and attribute-based strategies matching appears as a natural elicitation procedure (e.g., Payne et al. (1998)) and is widely used for certain or uncertain decision-making problems (see for instance, Keeney and Raiffa (1976) for utility assessment methods under uncertainty and Bana e Costa and Vansnick (1994), Edwards and Barron (1994) for utility assessment without uncertainty). Hence, matching is considered as a suitable tool for preference elicitation since it offers a good compromise between the effort required from the DM and the obtained preferential information.

Nevertheless, it is important to analyze empirically how DMs actually respond to matching questions, and if their answers conform to rationality principles expected by the preference models. For instance, suppose a DM assesses a matching between two alternatives a and b, providing a value on one criterion for alternative b; a reverse matching, normatively equivalent, can be created by integrating the answer to the description of b and letting the
DM adjust a performance value on the previously totally specified option $a$. The answers to a bi-matching (these two corresponding matching questions) should be consistent, i.e., isopreferences resulting from the “forward” and “backward” matching should be identical, or should not vary significantly. Previous experimental studies (Delquié (1997), Willemsen and Keren (2002), Willemsen and Keren (2003)) showed that DMs significantly deviate from such consistency in a bimatching. In this paper, we extend these previous results and study the effect of the multicriteria conflict, or trade-off size, on the asymmetry of judgment between forward and backward matching questions.

In pairwise comparisons, multicriteria conflict arises when each alternative outperforms the other one on some criterion, which imposes a trade-off. Comparing two alternatives can be difficult if their respective advantages are of high magnitude (Deparis et al. (2012)). We propose an experimental setting to investigate the following questions: Is the observed difference in adjustment increased by the multicriteria conflict in a bi-matching task? Do the observed effects depend on the direction of the matching (evaluating an advantage or a disadvantage in the matching). As some elicitation procedures use matching questions with pairs of alternatives involving a strong conflict between criteria (e.g. Bana e Costa and Vansnick (1994), Edwards and Barron (1994), Keeney and Raiffa (1976)), we aim at analyzing the consequences of using such alternatives in matching questions when eliciting the preferences of a DM.

The paper is organized as follows. The next section is devoted to a description of the matching procedure. In Section 3, we formulate hypotheses about the effect of multicriteria conflict on DMs’ answers. The experimental design is described in Section 4; the results are presented in Section 5 and discussed in Section 6. In the last section, we draw some conclusions and propose further research avenues.
2. Eliciting preferences using matching

2.1. Matching procedure

In a matching procedure the DM is faced with two alternatives \( a \) and \( b \) with all evaluations of \( a \) being known, and \( b \) being evaluated on all criteria except one. The DM has to fill in the missing evaluation such that she is indifferent between \( a \) and \( b \). In a bicriterion setting, a matching between two alternatives \( a=(a_1, a_2) \) and \( b=(b_1, ?) \) can be denoted as: \((a_1, a_2) \ I \ (b_1, ?)\). alternatives differing on two of the criteria and equivalent in all the others. In the following we will restrict our presentation to the bicriterion case.

In the above example, \( a \) is called the *stimulus alternative*, \( b \) the *adjusted alternative*. The criterion for which an evaluation is missing is called the *adjusted criterion* (criterion 2 in the above example). The non dominance between two alternatives implies that some values should not be possible for \( b_2 \): for instance in the case where the two criteria must be maximized, if \( a_1 > b_1 \), the DM will choose \( b_2 \) greater than \( a_2 \).

Matching questions can be distinguished according to the direction of the matching (Figure 1):

- **matching up**: The DM estimates an advantage with respect to the adjusted criterion. When \( a_1 > b_1 \) (as in Figure 1, Matching up), adjusting on criterion 2 comes down to evaluate the proper advantage of \( b \) over \( a \) on criterion 2 \((b_2 > a_2)\), in order to compensate for the inferior evaluation on criterion 1.

- **matching down**: The DM estimates a disadvantage with respect to the adjusted criterion. When \( a_2 < b_2 \), (as in Figure 1, Matching down) adjusting on criterion 1 comes down to evaluate the proper disadvantage of \( b \) compared to \( a \) on criterion 1 \((b_1 < a_1)\), in order to compensate for the superior evaluation on criterion 1.

Another important characteristic of a matching question is related to the intensity of the multicriteria conflict involved. We define the *conflict* of the matching \((a_1, a_2) \ I \ (b_1, ?)\) as the difference of evaluation between the
two matched alternatives on the criterion for which both of them have an evaluation, i.e $|a_1 - b_1|$. Notice that this notion is relative in the sense that it only allows to compare different matchings on the same attribute.

Lastly, considering a reference point $R^*$, the valence of a criterion in a given matching question indicates the relative position of the evaluations of $a$ and $b$ on this criterion with respect to $R^*$. The valence of a criterion is

- **positive** if $a$ and $b$ have better evaluations on this criterion than $R^*$,
- **negative** if $a$ and $b$ have worse evaluations on this criterion than the reference point,
- **mixed** if one alternative has a better evaluation and the other alternative a worse evaluation than $R^*$.

2.2. Bimatching

A bimatching is composed by a sequence of two consecutive matchings that we call forward matching and backward matching.

The answer obtained in the forward matching is used as a stimulus in the backward matching and one stimulus evaluation of the forward matching is missing in the backward matching. Table 1 presents an example of
bimatching where \( A = (a_1, a_2) \) is the stimulus alternative in the forward matching, \( B = (b_1, b_2) \) is the adjusted alternative in the forward matching but the stimulus alternative in the backward matching, \( A' = (a_1, a_2) \) is the adjusted alternative in the backward matching.

<table>
<thead>
<tr>
<th>Task</th>
<th>Alternative a</th>
<th>Alternative b</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>forward matching</td>
<td>( (a_1, a_2) )</td>
<td>( (b_1, ?) )</td>
<td>( b_2 )</td>
</tr>
<tr>
<td>backward matching</td>
<td>( (a_1, ?) )</td>
<td>( (b_1, b_2) )</td>
<td>( a_2 )</td>
</tr>
</tbody>
</table>

Table 1: An example of bimatching.

Consider the bimatching illustrated on Figure 2 (left), where the criteria are to be maximized. According to procedural invariance (Tversky and Kahneman (1986)), one should elicit the same trade-off during the two matchings. Hence, the value elicited during the backward matching \( (a'_1) \) should be the same as the stimulus value displayed during the forward matching \( (a_1) \). It is interesting to analyze, for different combinations of matchings, the value of \( |a'_1 - a_1| \) which is expected to be null according to the normative theory, or at least close to zero on average if we consider some noise. Let us consider situations where \( |a'_1 - a_1| \neq 0 \). In the up-up bimatching described in Figure 2 (left), \( a'_1 \in [b_1, a_1[ \). Similarly, in the down-down bimatching described in Figure 2, \( a'_1 \in ]a_1, b_1] \). In both situations, \( a' \) deviates from \( a \), on the adjusted criterion, towards \( b \). We call this situation an underadjustment. Conversely, on the adjusted criterion, \( A' \) could deviate from \( a \) away from \( b \) (e.g. \( a'_1 > a_1 \) on Figure 2, left). We call this situation an overadjustment. In order to measure both the intensity of the asymmetry and its direction (underadjustment or overadjustment), we define \( \delta = a_1 - a'_1 \) for an up-up bimatching, and \( \delta = a'_1 - a_1 \) for a down-down bimatching. With this definition, \( \delta > 0 \) when there is underadjustment and \( \delta < 0 \) when there is overadjustment. Different types of bimatching exist according to the direction of each matching (up-up, up-down, down-up, down-down) and according to the adjusted criterion. Figure 2 depicts up-up and down-down bimatchings when \( \delta \) is evaluated on criterion 1. Note that if the direction is kept constant for both matchings of a bimatching, then the adjusted criterion must necessarily change between the two matchings.
2.3. Literature on matching questions

The elicitation of indifference curves has been studied for several decades and there exists a large literature (e.g. Slovic (1975)), but the bimatching procedure was first introduced by Hershey and Schoemaker (1985) in order to analyse the procedural invariance in the context of decision under risk. They compared two ways of eliciting utility functions: the DM must either specify the certainty equivalent of a lottery or specify the probability of gain for an equivalent lottery. Delquié has adapted this notion to the context of certain multicriteria decision making (see Delquié (1993) and Delquié (1997)).

A number of studies shows that the trade-off elicited during a matching procedure with two criteria depends on the adjusted criteria (Fischer and Hawkins (1993), Slovic et al. (1990)). Using experiments on lotteries and also with certain alternatives, Delquié observed underadjustment during the backward matching. The estimated value is between the benchmark value and the target as in Figure 2.

It is interesting to note that the effects observed during the bimatching procedure are more sophisticated than the effects observed when two different elicitation procedures are used, since in bimatching the same procedure is used, the only modifications concerning the configurations of adjusted criteria and direction.
By rearranging the results obtained by Delquié Delquié (1993) depending on the direction of each matching, Willemsen and Keren (2003) showed that a significant asymmetry is more frequently observed during a down-down matching than an up-up matching. They also observed an effect of valence: if the adjusted criterion has a negative valence (i.e. it necessarily has a bad evaluation with respect to the reference alternative) one observes an asymmetry between the two tasks: the adjustment is weaker during the down matching.

3. Hypotheses

As mentioned in Subsection 2.3, previous experiments on the bimatching procedure have shown the influence of two factors: direction and valence on the answers to bimatching. Our experiment aims to improve these results by analysing the impact of multicriteria conflict on preferences elicited using a bimatching task. We also want to study the interaction of conflict with the two variables previously mentioned, direction of the matching and valence of the criteria. Our hypotheses are the following:

Hypothesis 3.1. In a bimatching task, the observed difference in adjustment increases with the multicriteria conflict.

During the forward matching, the DM is asked to match $b$ (adjusted alternative) with $a$ (stimulus) by specifying the evaluation of $b$ on the adjusted criterion. During the backward matching, (s)he is asked to match $a$ with $b$, by specifying the evaluation of $a$ on the adjusted criterion. The difference in adjustment refers to the difference in evaluation between $a$ and $b$ on the criterion adjusted during the backward matching (i.e. $\delta$ on Figure 2).

Our hypothesis states that increasing the difference in evaluation between $a$ and $b$ (on the specified criterion) during the forward matching increases the difference in adjustment. In other words, a larger conflict in a bimatching task reduces the consistency between the answers to each matching.

Hypothesis 3.2. The difference in adjustment observed during a down-down bimatching is higher than during an up-up bimatching. Moreover, this effect of the direction of matching increases with multicriteria conflict.
Willemsen and Keren (2003) observed that the difference in adjustment was significantly higher in a down-down bimatching task than during an up-up bimatching task. We aim at replicating such empirical finding. In addition, we refine the hypothesis, stating that the effect of direction increases with conflict, i.e with the difference in evaluation between $a$ and $b$.

**Hypothesis 3.3.** *The reference point used by the DM allows to define a valence for all criteria. The difference in adjustment observed during a bimatching is larger when the adjusted criterion is of negative valence than when it is of positive valence. This effect of valence increases with multicriteria conflict.*

The valence of the criteria involved in a bimatching is relative to a reference point $R^*$ (status quo alternative). In bimatching tasks involving criteria with positive valence (respectively negative valence) $a$, $b$, $a$ have better (respectively worse) evaluations than $R^*$. Hence, in bimatchings involving criteria with positive valence (advantageous bimatchings) the evaluation of $a$, $b$, $a$ are encoded as gains with respect to $R^*$ whereas in those involving criteria with negative valence (disadvantageous bimatchings), evaluations are encoded as losses.

Willemsen and Keren (2003) already reported that the difference in adjustment is greater in disadvantageous bimatchings than in advantageous ones. We complement this hypothesis stating that such effect of valence is increased when the difference of evaluation between $a$ and $b$ increases.

4. Method

In this section, we describe the behavioral experiment designed to test the hypotheses described in the previous section.

4.1. Participants

Twenty-nine graduate students (14 boys and 15 girls) in engineering at Ecole Centrale Paris, France, individually completed this experiment in re-
turn for 15€. They were native French speakers. The experimental instructions (available in appendix) were read to them by the experimenter. The stimuli and additional instructions were delivered through the E-Prime software program, on a computer screen. Subjects had to consider apartments to rent. Students in their final year have been chosen because they have to achieve a training course abroad. As a result, judging apartments was a meaningful task for them.

4.2. Stimuli

The basic task is a matching between two apartments represented only by the associated rent and the commuting time to city center, and equivalent on any other criteria. Figure 3 represents an example of stimulus. The evaluations of the two apartments are displayed in light gray boxes where the rent is indicated as an amount in Euros and the commuting time is indicated in minutes. The instructions inform the participants that their workplace as well as three features they have designated as important during a preliminary survey (e.g. theater, sports center, library) are located in the city center.

Evaluations on both criteria are available for Apartment A and displayed in the left box, while one evaluation is missing for Apartment B, represented in the right box. Using the keyboard, the participant has to indicate the missing evaluation for B, such that A and B provide her with the same satisfaction. The value the participant must provide is a rent during forward matchings, and a commuting time during backward matchings.

An apartment playing the role of reference option is associated to each participant in order to provide a valence to the criteria. A dark grey box, displayed on the lower part of the screen, represents this apartment. For the first group of participants this apartment dominates apartments A and B, hence the criteria both have a negative valence during all matchings. On the contrary, the reference apartment for the second group of participants is dominated by apartments A and B. Hence the criteria both have a positive valence during all matchings.
Figure 3: Example of stimulus during the main task.

4.3. Procedure

4.3.1. Context.

At the beginning, participants were informed of the context in which they were to choose their accommodation: their internship was taking part in a big city, remote from their family. The compensation of their training was 1600€/month, and they had to cover all common costs, including rent, using this amount. After receiving instructions, participants were asked to fill in an on-screen survey about their habits and tastes regarding accommodation. The aim of this survey was to acclimate the participants to the context of evaluating an apartment based on its rent and commuting time to the city center. For instance, the participants had to choose their three favorite places to have close to home (like a museum or a sports centre). Then we elicited rent levels: we asked them to determine seven monetary amounts typical of
their feeling about rent, from very unattractive to very attractive. In the same way, seven commuting time levels were elicited (Table 2 represents as an example the commuting time levels of one participant).

The participants were then informed that an apartment had already been pre-reserved for them. However a difficulty related to this apartment was leading them to look after another one. For Group 1 (negative valence), this apartment had very good evaluations on both criteria but they were not sure of getting it; for Group 2 (positive valence) the situation was opposite: the apartment was immediately available but had very poor evaluations on both criteria. Given the poor evaluations of the reference apartment, it was more interesting for them to keep searching for apartments. In Group 1 (Group 2), participants were informed that all apartments subsequently displayed would have worse (better) evaluations on both rent and commuting time.

Before starting the main task, participants received a detailed explanation about the matching process and had to answer to three test matchings.

4.3.2. Main Task.

The experimental task was a matching. Each participant answered to 64 matching questions, corresponding to 32 bimatchings.

In the problem under study, both criteria (Rent and Commuting time) are minimized by the subjects. On all figures, the scales associated to the X-axis and Y-axis are decreasing in minutes and euros respectively: Commuting time decreases when moving to the right (marginal value for Commuting time increases). Rent decreases when moving upwards (marginal value for Rent increases).

4.4. Main Task Design

Our experimental hypotheses involve three notions: direction, valence and conflict involved in a bimatching. To control the effect of the direction of matching we use bimatchings with a constant direction (up-up or down-down) in order to have similar effects in both matchings. Constant direction imposes that the DM does not adjust the same criterion during the forward matching
and the backward matching. We chose Rent as the adjusted criterion for all forward matchings and Commuting time for all backward matchings.

We present in the following the variables of our experiment (the first three are controlled variables and the last one is the observed variable).

**Direction** For each participant, half bimatchings are up-up and the other half are down-down. In other words, the participant either makes improving judgements at both steps or makes degrading judgements at both steps, depending on the bimatching.

**Conflict** It is defined by the forward matching. It only depends on the difference of evaluation on Commuting time between the two alternatives. We defined four levels of conflict by using the individual scale obtained at the beginning of the experiment. We ignored the two extreme levels of the scale (worst location and ideal location). For instance, for the subject whose preference scale on Commuting time are represented in Table 2, there is only one possible matching with conflict level four (very attractive vs. very unattractive, i.e. 15 minutes vs. 90 minutes), two possible matchings with conflict level three (15 minutes vs. 60 minutes and 25 minutes vs. 90 minutes), etc. For a given level of conflict, the displayed matching was chosen randomly among the different possible matchings.

**Valence** For group 1 (group 2), the two criteria had a negative (positive) valence.

**The asymmetry of adjustment** ($\delta$) It is the dependent variable. It is defined as in Subsection 2.2, knowing that criteria have to be minimized in this experiment. $\delta$ is computed on the criterion Commuting time as:

$$\delta = \begin{cases} 
\text{time}_{A'} - \text{time}_A & \text{for down-down bimatchings} \\
\text{time}_A - \text{time}_{A'} & \text{for up-up bimatchings}
\end{cases}$$

where $A$ is the stimulus alternative used during the forward matching and $A'$ is the response alternative in the backward matching. $\delta > 0$ means the bimatching resulted in an underadjustment, and $\delta < 0$ means the bimatching resulted in an overadjustment.
Table 2: An example of commuting time scale obtained for a participant

<table>
<thead>
<tr>
<th>description</th>
<th>time (min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ideal</td>
<td>5</td>
</tr>
<tr>
<td>very attractive</td>
<td>15</td>
</tr>
<tr>
<td>attractive</td>
<td>25</td>
</tr>
<tr>
<td>neutral</td>
<td>30</td>
</tr>
<tr>
<td>unattractive</td>
<td>60</td>
</tr>
<tr>
<td>very unattractive</td>
<td>90</td>
</tr>
<tr>
<td>worst</td>
<td>120</td>
</tr>
</tbody>
</table>

In order to account for the difference in scale range between participants, on the criterion Commuting time (one participant may never consider commuting times higher than 60 min, while another may consider times as high as 120 min), the data has to be normalized according to the respective scale range of each participant. We define $\delta_{\text{norm}} = \delta/(t_{\text{max}}(i) - t_{\text{min}}(i))$, where $t_{\text{max}}(i)$ is the commuting time considered as very unattractive by participant $i$, and $t_{\text{min}}(i)$ is the commuting time she considered as very attractive. During the elicitation of Commuting time values to anchor the scale at the beginning of the experiment, participants also assessed a worst commuting time and an ideal commuting time (respectively 1st and 7th echelons on the scale), but these values were not used to create stimuli during the experiment, so they were also ignored for normalization.

Each participant responded to 8 bimatchings for each level of conflict (4 up-up bimatchings and 4 down-down bimatchings). The order of bimatchings was random. Each matching was randomly drawn from a list containing 32 forward matchings to start. After each answer to a forward matching the corresponding backward matching was added to the list.

5. Results

We analyzed the effect of Multicriteria conflict, Direction and Valence on the value of $\delta_{\text{norm}}$. In total, there were 32 pairs of answers per participant, with 8 per level of Multicriteria conflict (4 up-up bimatchings and 4
down-down bimatchings). Hence, for a given combination of Direction and Multicriteria conflict, each participant answered to four bimatchings. When an analysis of variance is carried out, the mean of the four resulting $\delta_{\text{norm}}$ values was used as input data. Final data was made of 480 couples of answers for Group 1 and 448 for Group 2.

Of the 928 pairs of answers to bimatchings, only 114 were perfectly symmetrical (i.e. such that $\delta = 0$). A vast majority of the remaining pairs displayed underadjustment: 725 underadjustments (356 during up bimatchings, 369 during down bimatchings), compared to 89 overadjustments (40 during up bimatchings, 49 during down bimatchings). In both up-up bimatchings and down-down bimatchings, the answer to the backward matching does not correspond to the stimulus of the forward matching. Rather, this answer deviates in a way that makes the indifferent alternatives after the backward matching closer on Commuting time than the indifferent alternatives after the forward matching.

A three-way ANOVA showed no significant effect of Valence (average $\delta_{\text{norm}} = 0.19$ for positive valence vs. 0.21 for negative valence, $p = 0.30$), and no significant interaction of Valence with any other factor or combination of factors (interaction with Multicriteria conflict: $p = .25$, with Direction: $p = .83$, with Direction $\times$ Multicriteria conflict: $p = .80$). In order to test the effect of the within-subject factors with more precision, we carried a two-way ANOVA on the data after transformation in order to cancel inter-subject differences (the $\delta_{\text{norm}}$ were all translated to the mean $\delta_{\text{norm}}$ across all participants, irrespective of Valence).

Figure 6 illustrates the data distributed along the within-subject factors. One can observe that Direction and Multicriteria conflict both influence the values of $\delta_{\text{norm}}$. The analysis of variance shows that Direction exerts a significant effect ($F(1,28) = 15.9$, $p < .001$). When answering to down bimatchings, the corresponding asymmetry between forward and backward matching represents 24% of the scale range on average, while it represents 16% of the scale range for up bimatchings. The asymmetry between forward and backward matchings also differed significantly across the four levels of Multicrite-
ria conflict ($F(3, 84) = 193.5, p < .001$). A post-hoc analysis (Tukey’s HSD) shows that any increase in the level of Multicriteria conflict has a significant effect on the $\delta_{\text{norm}}$. The more dissimilar are the alternatives on Commuting time during a forward matching, the larger the deviation between the answer and the stimulus during the backward matching. Finally, a significant interaction exists between Multicriteria conflict and Direction ($F(3, 84) = 7.4, p < .001$). For both directions of bimatching, Multicriteria conflict increases the resulting asymmetry, but the effect is stronger during up bimatchings. Tukey post-hoc comparisons indicate that $\delta_{\text{norm}}$ differed significantly between up bimatchings and down bimatchings at Multicriteria conflict level 3 (0.21 up vs. 0.29 down, $p = .002$) and at level 4 (0.29 up vs. 0.44 down, $p = .001$). By contrast, the difference was not significant at Multicriteria conflict level 1 and level 2.
6. Discussion

6.1. Interpretation

First, our experiment replicates the main result of Delquié Delquié (1993) on bimatchings: there is an asymmetry of adjustment between the forward matching and the backward matching and the answers of the backward matching deviate from the values of the forward matching in the direction of the stimulus.

In Willemsen and Keren (2003), Willemsen and Keren observed a significant deviation only in the case of down-down bimatchings while our results also show a deviation for up-up bimatchings. This difference may be related to the fact that we control the conflict level and assign some bimatchings with a large multicriteria conflict. It is possible that the bimatchings of Willemsen and Keren (2003) only involved conflict levels large enough to induce significant deviations for up-up bimatchings only. If we consider our results for level
Figure 6: Effect of Multicriteria conflict and Direction on the asymmetry of adjustment between forward matching and backward matching, as a proportion of the scale range.

1 conflict only, we also obtain a significantly non-null asymmetry for down bimatchings only (down: $t(28) = 4.68, p < .001$, up: $t(28) = 1.87, p = 0.07$).

Our experiment assesses the effect of multicriteria conflict on the asymmetry of adjustment. This effect is stronger for down-down bimatchings. For bimatchings with the highest level of conflict, the amplitude of asymmetry (29 minutes on average for down-down bimatchings and 19 minutes for up-up bimatchings) is considerable. The difference between the two end-levels of scale is about one hour on average. Hence the asymmetry takes up a large part of the scale (see Figure 7). As a comparison, the asymmetry observed in Experiment 1 of Willemsen and Keren (2003) represents 14.7% of the scale they consider.

No significant effect of valence is observed. However, we think that participants did not consider the provided reference alternative when matching, which prevented from really measuring any potential effect. Hence, we cannot conclude on the effect of valence on the asymmetry of bimatching. Our experimental set-up did not guarantee that a given subject would consider
Figure 7: The effect of direction and conflict on the ratio asymmetry of adjustment/difference on commuting time between alternatives in the forward matching.

the provided reference apartment and not their idea of a typical reference apartments or the stimulus apartment as a reference point.

6.2. An explicative model of results: Anchoring on the stimulus alternative

In this section, we investigate if the experimental results can be explained by gain/loss asymmetry, assuming that participants considered the stimulus alternative as a reference point. The implicit use of the stimulus alternative as a reference point is plausible for several reasons: first, during a matching task, the stimulus alternative, against which the participant is matching, is entirely specified. Like the explicitly provided reference point, the stimulus alternative is easy to visualize. Second, it is the stimulus alternative that sets the overall value characterizing the matching task. The indifference curve corresponding to the matching task is identified by the stimulus alternative. Hence, it is natural for the DM to ask herself: “Relative to this alternative, how much am I ready to lose on Commuting time, in order to gain such amount on Rent?”, framing the matching as a gain/loss relative to
If the DM does consider the stimulus alternative as a reference point, gain/loss asymmetry in the encoding of marginal values intervenes in the adjustment made by the DM in order to answer to the matching task. On a given criterion, the loss of a certain quantity on evaluation has a stronger effect on marginal value variation than the gain of the same quantity (see Tversky and Kahneman (1991) and Highhouse and Johnson (1996)), which can be summed up as: “Losses loom larger than corresponding gains”. Graphically, gain/loss asymmetry can be represented by a marginal value curve which is convex for losses and concave for gains. If we decompose a bimatching, we can identify how gain/loss asymmetry should affect the judgements occurring at the various steps of the bimatching. In the following, we consider traded quantities of evaluation, i.e., quantities subtracted or added on an evaluation when moving from one alternative to the other. $\Delta t$ refers to a traded quantity of evaluation on the criterion Commuting time, in minutes, and $\Delta r$ refers to a traded quantity of evaluation on the criterion Rent, in euros. On all figures in this section, Rent and Commuting time are represented on axes decreasing in euros and minutes, respectively, so that maximizing marginal value is going in the up-right direction.

![Diagram](image)

**Figure 8:** Adjustment $-\Delta r$ during a down forward matching.
Let us consider a down-down bimatching. During the forward matching, \((A(t_A, r_A) \ I \ B(t_B, ?))\), see Figure 8), the DM observes that \(B\) outperforms \(A\) by \(+\Delta t\). This translates in a marginal value gain \(v(+\Delta t)\). To keep the overall value constant when moving from \(A\) to \(B\), the DM must compensate for that gain by a loss of same marginal value, on Rent. Hence, she must assess a disadvantage on Rent \(-\Delta r\) such that:

\[
|v(-\Delta r)| = v(+\Delta t)
\]

This equality represents the expected indifference at the end of the forward matching.

During the backward matching, the DM observes that \(A\) outperforms \(B\), on Rent, by a quantity \(\Delta r\) (see Figure 9) By design, this quantity is the one the DM chose during the forward matching. But at this step, the DM will consider this quantity as a gain. She must trade the marginal value gain \(v(+\Delta r)\) against a loss \(-\Delta' t\) on Commuting time, such that:

\[
|v(-\Delta' t)| = v(+\Delta r)
\]

This equality represents the expected indifference at the end of the backward matching.
How do $\Delta t$ and $\Delta't$ compare? The same quantity $\Delta r$ is involved during both matchings. However, its psychological value is not the same during the two steps. Gain/loss asymmetry implies the following inequality:

$$v(+\Delta r) < |v(-\Delta r)|$$

By replacing both terms of the inequality using 1 and 2: $|v(-\Delta't)| < v(+\Delta t)$. If we now apply gain/loss asymmetry to the Commuting time criterion, $v(+\Delta't) < |v(-\Delta't)|$. Finally, by transitivity: $v(+\Delta't) < v(+\Delta t)$. And since the marginal value function for Commuting time is monotonic strictly increasing:

$$+\Delta't < +\Delta t$$

This inequality follows from gain/loss asymmetry, applied to each criterion, and assuming that the DM trades equal amounts of marginal value when matching two alternatives. Our demonstration assumes that the DM uses the stimulus alternative as a reference point and predicts an underadjustment: the difference on Commuting time between the matched alternatives is smaller after the backward matching than the stimulus difference existing during the forward matching.

6.3. Implications for Preference Elicitation

The results of our experimentation can be used to formulate some recommendations regarding the elicitation process. There is a significant variation in the elicited trade-offs depending on the choice of the adjusted criterion. This asymmetry increases in a drastic fashion with multicriteria conflict. This observation should lead any analyst to favour matching tasks imposing a low level of multicriteria conflict. Yet, many elicitation methods proposed for utility functions make use of examples with extreme levels of conflict. For example, the swing weights procedure, used in the SMARTS and SMARTER methods presented in Edwards and Barron (1994), involves comparisons of fictitious alternatives with the best evaluation on some attributes and the worst on all other attributes. In the first step, which is aimed at ranking attributes, the DM has to express his/her preferences about the set of all these extremely specialized alternatives. In the second step, he/she has to match
a specialized alternative with another alternative having the worst evaluation on all attributes except one. By design, such tasks involve high, and possibly the highest, bicriteria conflict. Such a difficulty also arises in the method proposed by Keeney and Raiffa (1976) to elicit weights. Even if we consider a decision problem where the fictitious specialized alternatives are realistic, such a degree of conflict should make the analyst careful about how to interpret the preferential information that he/she obtains. For this type of approaches, it can be more careful to propose a bimatching, preferably hidden, and to use the average trade-off, if an analysis of sensitivity is impossible. Other methods avoid a large conflict, by involving carefully selected fictitious alternatives, such as MACBETH (Bana e Costa and Vansnick (1994)). While this method involves comparisons between fictitious alternatives differing on at most two criteria, the alternatives do not have the maximum or minimum evaluation on each attribute. Instead, the performances on each attribute are kept at a reference level chosen by the analyst (often a neutral and an attractive level) assigning comparisons with a controlled level of conflict. Another way to avoid high multicriteria conflict is to consider real alternatives that the DM can easily compare (e.g. Jacquet-Lagreze and Siskos (1982) and Greco et al. (2008)).

The results of our experiment can also be interpreted as revealing a type of intransitivity of indifference. During successive forward matchings and backward matchings, we obtained affirmations of the following kind: $aIb$ and $bIa'$. And yet alternatives $a$ and $a'$ cannot be indifferent since they share the same evaluation on one criterion and the difference on the other criterion may be very large (such as half of the scale). This observed intransitivity differs in nature from the classical interpretation of intransitivity of indifference related to the cumulation of small indifference thresholds (as in the famous coffee example by Luce (1956) used for semiorders). Here instead, the differences in evaluation may be very large. We think that in our experiment the observed intransitivity has more to do with the difference in strategies used during the forward matching and the backward matching.
7. Conclusions

In this paper, we are interested with the effect of multicriteria conflict on the expression of preferences through the matching procedure. Multicriteria conflict refers to the difference in evaluation, along criteria, between two non mutually-dominating alternatives. In the context of matching, in a bicriteria decision, it is defined as the difference in evaluation between the matched alternatives on the criterion where they are both specified. Using bi-matchings, where the answer to a first matching is used as a stimulus in a second, reverse matching, we observed that expressed preferences become inconsistent when multicriteria conflict is increased. An asymmetry arises between answers to the forward and backward matching, in a systematic direction: the compensation rate is more favourable to the adjusted criterion. Direction of the matching also affects the asymmetry of answers: degrading bi-matchings lead to a higher difference between compensation rates than improving bi-matchings. The effect of direction is increased by multicriteria conflict, resulting in differences of evaluation that can reach half of the scale, between the benchmark (stimulus of the first matching) and the answer to the reverse matching.

One of the implicit assumptions behind elicitation through matching is that isopreference curves faithfully describe the multicriteria preferences of the DM. Yet, our experimental results show that what DMs express through matching judgements is not as clearly defined as isopreference curves. We observed systematic underadjustments, when matching apartments on commuting time, compared to matching on rent. The size of the underadjustment was increased both by multicriteria conflict, and by asking degrading rather than improving judgements. The asymmetry of the object we observed through matching contradicts one of the basic features of isopreferences. Further research is needed to better understand the nature of preferences obtained through matching.

Acknowledgement

The authors gratefully acknowledge Walid Oukachbi for his support in conducting the research.
References


