

Cumulative dominance in multi-attribute choice: benefits and limits

Konstantinos V. Katsikopoulos · Martin Egozcue ·
Luis Fuentes Garcia

Received: 4 November 2013 / Accepted: 30 April 2014 / Published online: 27 May 2014
© Springer-Verlag Berlin Heidelberg and EURO - The Association of European Operational Research Societies 2014

Abstract Guided decision processes improve decisions by employing personalized rules that are simple and perform well. In multi-attribute choice, simple rules such as deterministic elimination by aspects can achieve maximum utility because they choose a cumulative-dominant alternative when one exists and because of the combination of two results: (i) a cumulative-dominant alternative has maximum utility when the utility function is linear and (ii) a cumulative-dominant alternative exists frequently in the majority of choice problems when the attributes are binary. We test the limits of (i) and (ii) by relaxing their assumptions about the utility function and the attributes: First, for a linear utility function that also includes multiplicative attribute interactions, we analytically show that a cumulative-dominant alternative has maximum utility when attributes are binary. Whereas this is not necessarily true for continuous attributes, we empirically find that it is essentially always true for a range of probability distributions of the attributes and parameterizations of the utility function. Second, we empirically find that when attributes are continuous, a cumulative-dominant alternative exists frequently only in the minority of choice problems. In sum, cumulative dominance legitimates guiding

K. V. Katsikopoulos (✉)
Center for Adaptive Behavior and Cognition (ABC), Max Planck Institute for Human Development,
Lentzeallee 94, 14195 Berlin, Germany
e-mail: katsikop@mpib-berlin.mpg.de

M. Egozcue
Department of Economics, Facultad de Ciencias Sociales, Universidad de la Republica,
11200 Montevideo, Uruguay
e-mail: megozcue@decon.edu.uy

L. F. Garcia
Departamento de Metodos Matematicos e de Representacion, Universidade da Coruna,
15071 La Coruña, Spain
e-mail: luis.fuentes@udc.es

decision processes by simple rules when it exists but its existence should not be taken for granted.

Keywords Decision analysis · Guided decision processes · Multi-attribute choice · Cumulative dominance · Simple heuristics · Elimination by aspects

Mathematics subject classification 60 · 90

1 Introduction: simplifying multi-attribute choice

Baucells and Sarin (2013) argue for guiding decision processes by employing personalized decision rules that “require only a few meaningful inputs from the decision maker” and “have strong prescriptive properties” (p. 30). In other words, the rules should be simple and perform well. For example, such simple rules are deterministic elimination by aspects (Tversky 1972) and equal weighting of attributes (Dawes and Corrigan 1974), which perform surprisingly well in multi-attribute choice (Hogarth and Karelaia 2005a, b; Katsikopoulos 2013).

A condition that facilitates the performance of simple rules in multi-attribute choice is cumulative dominance (Kirkwood and Sarin 1985): Say that there are three attributes and the utility function is $U(X) = 3a_1(X) + 2a_2(X) + a_3(X)$. There are two alternatives, A with attribute values $a_1(A) = 1$, $a_2(A) = 0$ and $a_3(A) = 1$ and B with attribute values $a_1(B) = 0$, $a_2(B) = 1$ and $a_3(B) = 1$. Then, A cumulatively dominates B because $a_1(A) \geq a_1(B)$, $a_1(A) + a_2(A) \geq a_1(B) + a_2(B)$ and $a_1(A) + a_2(A) + a_3(A) \geq a_1(B) + a_2(B) + a_3(B)$. The cumulative-dominant alternative A also maximizes utility since $U(A) = 4$ and $U(B) = 3$. Note that A is chosen by deterministic elimination by aspects (Hogarth and Karelaia 2005a, b) which, for two alternatives, orders attributes according to the magnitude of their weight and chooses the alternative with the higher value on the first attribute that has different values on the two alternatives (in the example, this is attribute a_1). In sum, one can make utility-maximizing choices by knowing only the order of attribute weights and not their exact magnitudes.

The result that cumulative-dominant alternatives maximize utility is not peculiar to this example but holds for all linear utility functions (Baucells et al. 2008). Furthermore, some simple rules such as deterministic elimination by aspects, always choose a cumulative-dominant alternative when one exists (Baucells et al. 2008) and cumulative-dominant alternatives exist frequently in the majority of choice problems with binary attributes (Şimşek 2013).

In sum, cumulative dominance seems to legitimate guiding decision processes by employing simple rules. But how robust is this result? What if the utility function is not linear or if the attributes are not binary? These are the questions we ask and answer here.

The remainder of this article is organized as follows: in the next section, we present the previous work that has demonstrated some benefits of cumulative dominance. Then we test, via mathematical analysis and computer simulation, the

limits of cumulative dominance by relaxing previous assumptions about the utility function and the attributes. The last section concludes.

2 Cumulative dominance: known benefits

The goal is to choose one out of many alternatives A, B, C, \dots , so as to maximize a function that maps alternatives to a numerical measure of their quality. This function can be interpreted as the decision maker's subjective utility function (Keeney and Raiffa 1976).

Each alternative A is described by n ($n \geq 2$) attributes, a_1, \dots, a_n . That is, $A = (0.8, 0.2, 0.5)$ means that $a_1(A) = 0.8$, $a_2(A) = 0.2$ and $a_3(A) = 0.5$. When the attributes are *continuous*, they can take any value in the interval $[0, 1]$. When the attributes are *binary*, they equal either 0 or 1. There exist k ($k \geq 2$) alternatives in the consideration set.

In Baucells et al. (2008), the true utility function, which is unknown to the decision maker and the decision analyst, is assumed to be a linear function of the attributes:

$$U(X) = \sum_{i=1}^n w_i a_i(X); \quad w_1 \geq w_2 \geq \dots > w_n \geq 0. \quad (1)$$

The utility function described in (1) is a special case of Keeney and Raiffa's additive utility function where each single-attribute utility function equals the identity function (1976, Section 6.4). Theoretically, sufficient conditions for the existence of an additive utility function are reviewed in (Keeney and Raiffa 1976, Theorem 6.3).

Baucells et al. (2008) assume that the utilities in (1) are used for ranking alternatives. We make the same assumption here.

Alternative A *cumulative-dominates* alternative B (Kirkwood and Sarin 1985) when the following holds:

$$\sum_{i=1}^j a_i(A) \geq \sum_{i=1}^j a_i(B); \quad j = 1, 2, \dots, n. \quad (2)$$

Note that (2) assumes an order of attributes. In what follows we fix this order and assume that it is the same with the order of the weights in (1).

We say that an alternative A is cumulative-dominant in the consideration set if it cumulative-dominates all other $k - 1$ alternatives.

A cumulative-dominant alternative has maximum utility when the utility function is linear. In the Introduction, we saw an example of this for binary attributes. It is also true for continuous attributes. For example, assume $U(X) = 3a_1(X) + 2a_2(X) + a_3(X)$ and consider two alternatives A and B with $a_1(A) = 0.8$, $a_2(A) = 0.2$ and $a_3(A) = 0.5$, and $a_1(B) = 0$, $a_2(B) = 1$ and $a_3(B) = 0.4$. It holds that A cumulatively dominates B and $U(A) = 3.3 > 2.4 = U(B)$. More generally, (Baucells et al. 2008, Theorem 3) analytically showed the following:

Result 1 *For the linear utility function (1), A cumulative-dominates B [see (2)] if and only if $U(A) \geq U(B)$ for all weights w_i .*

Note that Result 1, and all results in the present article, assume that the decision maker is properly calibrated: The decision maker, possibly with the help of the decision analyst, knows how to code the attributes a_i so that the weights w_i are positive and is able to order the w_i according to their magnitude. These assumptions are particularly plausible when the decision maker is familiar with the choice. In this sense, the scope of this article is decision making under risk, not uncertainty (Knight 1921).

Result 1 implies that if an alternative is cumulative-dominant in the consideration set, then it has maximum utility. The next question is how frequently do cumulative-dominant alternatives exist.

Baucells et al. (2008) investigated artificial choice problems where the values of the attributes were drawn from independent Bernoulli distributions or from simple generalizations of Bernoulli distributions that had attribute correlations. The number of attributes n and the number of alternatives k were varied from 2 to 10. When n increased, the probability of cumulative dominance decreased. The decrease, however, was not too rapid. For example, when attributes followed independent Bernoulli distributions and it was equally likely that an attribute equalled zero or one, the probability of cumulative dominance ranged from 97 to 68 % as n increased from 2 to 10 when there were $k = 2$ alternatives. For $k = 10$ alternatives, this probability ranged from 99 to 42 %. Combinations of n and k for which the probability of cumulative dominance fell below 50 % were rare and in most combinations this probability did not fall below 80 %. Furthermore, Şimşek (2013) investigated natural choice problems in 51 datasets with n ranging from 3 to 21 and $k = 2$. The probability of cumulative dominance ranged from 100 to 61 % with a mean of 87 %. We summarize these results informally as follows:

Result 2 *For binary attributes, a cumulative-dominant alternative exists frequently for the majority of combinations of the number of attributes n and the number of alternatives k .*

Based on Results 1 and 2, multi-attribute choice can be simplified if one has a method for choosing cumulative-dominant alternatives when such alternatives exist. Some heuristics, which are simple to apply and are often used spontaneously by people, indeed choose cumulative-dominant alternatives. For example, Hogarth and Karelaia (2005a, b) define a deterministic version of Tversky's (1972) elimination by aspects. In this heuristic the decision maker first inspects the value of all alternatives on attribute a_1 . If all alternatives have the same value on a_1 , this attribute is eliminated. If one alternative has a unique maximum value on a_1 , this alternative is chosen. Otherwise, all alternatives with a value smaller than the maximum on a_1 are eliminated and the process is repeated using attribute a_2 . The process continues in this way until an alternative is chosen. For example, the heuristic chooses $A = (0.8, 0.2, 0.5)$ over $B = (0, 1, 0.4)$ based on attribute a_1 . The performance of this and similar heuristics has been empirically found to compare favorably with statistical benchmarks such as linear and logistic regression,

Bayesian networks, neural networks and classification and regression trees (Czerlinski et al. 1999; Martignon and Hoffrage 2002; Chater et al. 2003; Hogarth and Karelaia 2005a; Brighton 2006; Martignon et al. 2008; Katsikopoulos et al. 2010, 2011; Şimşek 2013).

3 Cumulative dominance: testing the limits

Utility function with attribute interactions

In what follows, we assume the following utility function:

$$U(X) = \sum_{i=1}^n w_i a_i(X) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n w_{i,j} a_i(X) a_j(X) + \dots + w_{1,2,\dots,n} a_1(X) a_2(X) \dots a_n(X);$$

$$w_1 \geq w_2 \geq \dots \geq w_n \geq 0, w_{1,2} \geq w_{1,3} \geq \dots \geq w_{n-1,n} \geq 0, \dots, w_{1,2,\dots,n-1} \geq w_{1,2,\dots,n-2,n} \geq \dots \geq w_{2,3,\dots,n} \geq 0.$$

That is, (3) generalizes (2) by additionally including all multiplicative interactions among attributes (two-way, three-way, etc.). Note that the interaction weights are assumed to be positive and increasing in all subscripts. For example, for $n = 3$ attributes, the general form of the utility function is $U(X) = w_1 a_1(X) + w_2 a_2(X) + w_3 a_3(X) + w_{1,2} a_1(X) a_2(X) + w_{1,3} a_1(X) a_3(X) + w_{2,3} a_2(X) a_3(X) + w_{1,2,3} a_1(X) a_2(X) a_3(X)$ where $w_1 \geq w_2 \geq w_3 \geq 0, w_{1,2} \geq w_{1,3} \geq w_{2,3} \geq 0$ and $w_{1,2,3} \geq 0$ and a specific instance is $U(X) = 7a_1(X) + 6a_2(X) + 5a_3(X) + 4a_1(X) a_2(X) + 3a_1(X) a_3(X) + 2a_2(X) a_3(X) + a_1(X) a_2(X) a_3(X)$.

The utility function described in (3) is a special case of Keeney and Raiffa’s multi-linear utility function where each single-attribute utility function equals the identity function (1976, Section 6.4). Theoretically, sufficient conditions for the existence of a multi-linear utility function are reviewed in (Keeney and Raiffa 1976, Theorem 6.3). We provide empirical evidence that supports the particular assumptions made in (3).

First, in a meta-analysis of 113 datasets, Li et al. (2006) found that including multiplicative interactions in a linear model is crucial to fitting the data well since 11 % of the two-attribute, 7 % of the three-attribute and 3 % of the four-attribute interactions were statistically significant.

Second, there is evidence for the restrictions of (3) on the weights. Note that the assumption $w_i \geq 0$ can be made without loss of generality. Given this, it can also be assumed that $w_{i,j}, w_{i,j,k}, \dots, w_{1,2,\dots,n} \geq 0$ because of the empirical regularity of synergy which says that interaction weights tend to have the same sign with the corresponding single-attribute weights (in Li et al. 2006, this happened in 83 % of the statistically significant two-attribute interactions). Next, the assumption $w_i \geq w_{i+1}$ can be made without loss of generality. Given this, it can also be assumed that $w_{i,j}, w_{i,j,k}, \dots$ are decreasing in all subscripts because of an empirical regularity called inheritance (Wu and Hamada 2000; Li et al. 2006) which says that the magnitude of $w_{i,j}$ tends to depend only on the magnitudes of w_i and w_j and to increase when either of them increase (Chipman 1996).

Finally, we also tested how well (3) fits the data in 20 decision problems that have been used to evaluate the performance of various choice models (Czerlinski et al. 1999). In each of these problems, the task is to rank a number of alternatives according to their value on a criterion that can be interpreted as utility (e.g., in one problem the alternatives are counties in Minnesota and the criterion is their rent per acre and in another problem the alternatives are the contiguous states in the U.S. and the criterion is their average motor fuel consumption per person).

In each problem, fit was measured by the rank correlation (Kendall 1938) between the order of alternatives according to their utilities as estimated by (3) and the true order of alternatives according to their real criterion values. Specifically, the following procedure was used. First, it was determined which attributes to ignore by fitting (3) to all data in each problem and retaining only the attributes corresponding to the largest 25 % of the weights (when 50, 75 or 100 % were used the fit was essentially equal, so we chose 25 % to minimize the number of parameters). Second, the cross-validation method was used; Initially, a sample that contained 50 % of all alternatives, their attribute and criterion values, was used to estimate the weights that maximized the fit of (3) in this sample, and finally the fit of (3) with the estimated weights was measured in ranking the other 50 % of the alternatives. To average out random variation, this process was repeated 1,000 times. Across the 20 problems the average value of Kendall's τ was 0.43. And there was a 61 % probability that the alternative with the truly highest criterion value was the alternative with the maximum utility as estimated by (3).

3.1 Binary attributes: cumulative-dominant alternatives have maximum utility

We now ask if Result 1 can be extended for (3). To gain some intuition, consider the following examples. First, assume $U(X) = 7a_1(X) + 6a_2(X) + 5a_3(X) + 4a_1(X)a_2(X) + 3a_1(X)a_3(X) + 2a_2(X)a_3(X) + a_1(X)a_2(X)a_3(X)$. Take two alternatives A and B with $a_1(A) = 1$, $a_2(A) = 1$ and $a_3(A) = 0$, and $a_1(B) = 1$, $a_2(B) = 0$ and $a_3(B) = 1$. A cumulatively dominates B and $U(A) = 17 > 15 = U(B)$, so in this case Result 1 holds. But for the alternatives A and B with $a_1(A) = 1$, $a_2(A) = 0$ and $a_3(A) = 0$, and $a_1(B) = 0.8$, $a_2(B) = 0.1$ and $a_3(B) = 0.1$, Result 1 does not hold: A cumulatively dominates B but $U(A) = 7 < 7.28 = U(B)$.

So, Result 1 cannot be extended for a utility function with interactions (3) when the attributes are continuous. It can, however, be analytically shown that the result can be extended when the attributes are binary (for a proof, see the Appendix):

Result 3 *For the utility function with attribute interactions (3) and binary attributes, A cumulative-dominates B [see (2)] if and only if $U(A) \geq U(B)$ for all weights $w_i, w_{i, j}, \dots, w_{1, 2, \dots, n}$*

3.2 Continuous attributes: cumulative-dominant alternatives essentially always have maximum utility

In the absence of a theoretical guarantee that cumulative-dominant alternatives always maximize utility when the attributes are continuous, we investigated

empirically how frequently this is the case. In a computer simulation, we crossed two factors: (i) the probability distribution of the attributes and (ii) the parameterization of the utility function.

For (i), we used two symmetric and two skewed distributions. The symmetric distributions were the uniform distribution over the [0, 1] interval and the normal distribution with mean 0.5 and variance 1, truncated so that all values are in the [0, 1] interval. The skewed distributions were obtained by truncating the normal distribution with mean 0 and variance 1 and the normal distribution with mean 1 and variance 1. In all four cases, attribute values were drawn independently of each other.

For (ii), we used different sets of attribute and interaction weights. We used two types of weights that induce different compensations among the attributes (Hogarth and Karelaia 2005b; Katsikopoulos 2013). First, we used equal weights $w_i = w_j = \dots = w_1, 2, \dots, n = 1$ where compensation is maximum. Second, we used unequal weights where there is less compensation, with the largest weight being $w_1 = 2^n - 1$ and weights decreasing by 1 so that $w_2 = 2^n - 2, \dots, w_{2, 3, \dots, n} = 1$. Because the number of weights grows exponentially with n , we used all weights for n from 2 to 5 and only the single-attribute weights and the two-attribute interaction weights for n from 6 to 10.

The results are provided in Table 1. Each cell is the probability that the cumulative-dominant alternative (when one exists) has higher utility in a pair of alternatives ($k = 2$), for the corresponding combination of the number of attributes and the structure of weights. We repeated the process 10,000 times to average out random variation. Note that the results were always within 1 % for all four probability distributions and so we report the results only for the uniform distribution.

Because a cumulative-dominant alternative essentially always maximizes utility when there are two alternatives, it follows that a cumulative-dominant alternative essentially always maximizes utility for any number of alternatives as well.

3.3 Continuous attributes: a cumulative-dominant alternative exists frequently only in the minority of choice problems

We investigated empirically how frequently a cumulative-dominant alternative exists when attributes are continuous. In a computer simulation, we crossed three

Table 1 The probability of the cumulative-dominant alternative [when one exists, see (2)] having higher utility in a pair of alternatives ($k = 2$) as a function of the number of continuous attributes n (from 2 to 10) and the parameterization of the utility function (3) (equal or unequal weights; see text for details)

n	2	3	4	5	6	7	8	9	10
Equal weights	99	99	99	99	100	100	100	100	100
Unequal weights	100	100	100	100	100	100	100	100	100

The results were essentially identical for all four probability distributions used (see text for details) and so we report the results only for the uniform distribution

factors: (i) the probability distribution of the attributes (same as in the previous simulation), (ii) the number of attributes n (from 2 to 10) and (iii) the number of alternatives k (from 2 to 10). The results are provided in Table 2. Each cell is the probability that there exists a cumulative-dominant alternative for the corresponding combination of the number of attributes and the number of alternatives. We repeated the process 10,000 times to average out random variation. Note that the results were always within 2 % for all four probability distributions and so we report the results only for the uniform distribution.

How do the results of Table 2 compare with the results of Baucells et al. (2008) and Şimşek (2013) for binary attributes? A similarity is that in both cases the probability of cumulative dominance decreases with n . But the numbers are clearly lower when the attributes are continuous: for example, recall that for $k = 2$ and binary attributes, the probability of cumulative dominance ranged from 99 to 68 % (Baucells et al. 2008) and 100–61 % (Şimşek 2013). But in the top row of Table 2 we see that this probability ranges from 75 to 35 % when the attributes are continuous. The mean is 49 % which is comparable to the 58 % that Şimşek (2013) found across 51 natural datasets with continuous attributes (her range was from 94 to 11 %). Furthermore, only for 7 of the 81 combinations of n and k does the probability of cumulative dominance rise above 50 % in Table 2. And, for 44 of the 81 combinations, this probability does not rise above 20 %. We summarize these results informally as follows:

Result 4 *For continuous attributes, a cumulative-dominant alternative exists frequently for the minority of combinations of the number of attributes n and the number of alternatives k .*

In sum, the analytical and empirical results of the present section suggest that when attributes are binary, cumulative dominance exists frequently and leads to the good performance of simple heuristic rules; whereas when attributes are continuous, cumulative dominance exists less frequently and the performance of simple

Table 2 The probability that a cumulative-dominant alternative exists [see (2)] as a function of the number of continuous attributes n and the number of alternatives k

$n:$ k	2	3	4	5	6	7	8	9	10
2	75	63	55	50	45	42	39	37	35
3	64	48	40	35	30	26	24	21	19
4	58	41	31	27	22	20	17	15	13
5	52	35	27	21	18	16	13	12	10
6	49	31	23	18	15	12	11	9	8
7	45	28	20	16	12	10	9	7	6
8	43	25	18	14	11	9	8	6	6
9	40	23	16	12	10	8	7	5	5
10	38	22	15	11	9	7	6	5	4

The results were always within 2 % for all probability distributions used (see text for details) and so we report the results only for the uniform distribution

heuristic rules may be compromised. We illustrate our results on real-world choice problems from Şimşek (2013). That work reports summary statistics across 51 problems, whereas here we report results on two particular problems communicated to us by the author.

First, consider an individual who wants to buy a car that maximizes mileage per gallon of gas. In 1983, the Committee on Statistical Graphics of the American Statistical Association prepared a dataset including 398 cars built between 1970 and 1982, scored on the following continuous attributes: number of cylinders, engine displacement, horsepower, vehicle weight, time to accelerate from 0 to 60 mph, model year and origin. A cumulative-dominant alternative existed in 57 % of all comparisons between two cars but when attributes were binarized using a median split, a cumulative-dominant alternative existed in 89 % of all paired comparisons (for details on how the order of attributes used for checking cumulative dominance was determined, see Şimşek 2013). The cross-validated accuracy of the simple heuristic rule that chooses the cumulative-dominant alternative when one exists and guesses otherwise was 0.77 with continuous attributes and 0.84 with binary attributes.

A second example illustrates that the frequency of cumulative dominance and the performance of simple heuristics is not necessarily compromised when attributes are continuous. Consider an investor or a policy maker who wants to determine which of two anesthesiology departments requires fewer monthly man-hours to be maintained. Myers (1986) scored 12 American naval hospitals on the following attributes: number of surgical cases, eligible population and number of operating rooms (for details on the dataset, see Şimşek 2013). When attributes were binary, the frequency of cumulative dominance was 100 % and the cross-validated accuracy of the rule of choosing the cumulative-dominant alternative when one exists and guessing otherwise was 0.77, whereas when attributes were continuous, the frequency of cumulative dominance was 94 % and the cross-validated accuracy of the rule was 0.97.

4 Conclusions

Baucells and Sarin's (2013) *guided decision processes* are a decision-making approach which lies between the extremes of unaided intuition and demanding optimization. The idea is to employ personalized decision rules that are simple and also perform well as, for example deterministic elimination by aspects and equal weighting of attributes. Such heuristic rules are well studied in psychology (Tversky 1972; Dawes and Corrigan 1974; Payne et al. 1993; Gigerenzer et al. 1999) and first results have also been obtained in decision analysis (Martignon and Hoffrage 2002; Hogarth and Karelaia 2005b, 2006a, b; Katsikopoulos and Fasolo 2006; Katsikopoulos 2011, 2013).

The present work takes a general point of view by investigating cumulative dominance—a condition that facilitates the performance of simple rules in multi-attribute choice (Baucells et al. 2008; Carrasco and Baucells 2008; Şimşek 2013)—and delineating its boundary conditions. In sum, we found that when it exists,

cumulative dominance legitimizes guiding decision processes by employing simple rules but its existence should not be taken for granted. Our results are consistent with a growing body of work which has shown that simple heuristics are a good choice for making decisions, provided that the right heuristic is chosen for the right problem (Payne et al. 1993; Gigerenzer, Todd and the ABC research group 1999; Gigerenzer et al. 2011).

Acknowledgments We would like to thank Manel Baucells and Özgür Şimşek for their comments and helpful discussions.

Appendix

Result 3 For the utility function with attribute interactions (3) and binary attributes, A cumulatively dominates B [see (2)] if and only if $U(A) \geq U(B)$ for all weights $w_i, w_{i,j}, \dots, w_{1,2}, \dots, w_n$

Proof In order to simplify the exposition, we assume that $w_{i,j,k} = \dots = w_{1,2}, \dots, w_n = 0$. It will be clear that the argument also holds when this is not the case.

(\rightarrow) For any alternative X let $C(X) = \{i: a_i(X) = 1\}$. From (3), it follows that:

$$U(X) = \sum_{i \in C(X)} w_i + \sum_{i \in C(X)} \sum_{j \in C(X): j > i} w_{i,j}. \tag{4}$$

Because A cumulatively dominates B, (2) implies $\sum_{i=1}^n a_i(A) \geq \sum_{i=1}^n a_i(B)$. Because $w_i, w_{i,j} \geq 0$ for all i and j , it suffices to show that $U(A) \geq U(B)$ when $\sum_{i=1}^n a_i(A) = \sum_{i=1}^n a_i(B)$. Note that the latter, together with (2), implies the following:

$$\text{For any } i_A \in C(A), \text{ there exists an } i_B \in C(B) \text{ such that } i_A \leq i_B. \tag{5}$$

By combining (5) and the assumption of (3), $w_i \geq w_{i+1} \geq 0$ for all i , it follows that:

$$\sum_{i \in C(A)} w_i \geq \sum_{i \in C(B)} w_i. \tag{6}$$

By combining (5) and the assumption of (3), $w_{i,j} \geq w_{k,l} \geq 0$ for all $i \leq k$ and $j \leq l$, the following holds:

$$\sum_{i \in C(A)} \sum_{j \in C(A): j > i} w_{i,j} \geq \sum_{i \in C(B)} \sum_{j \in C(B): j > i} w_{i,j}. \tag{7}$$

Combining (4), (6) and (7) implies $U(A) \geq U(B)$ for all w_i and $w_{i,j}$.

(\leftarrow) Because $U(A) \geq U(B)$ holds for all weights $w_i, w_{i,j}$, it also holds for $w_i = 1/j^*$ for all $i = 1, 2, \dots, j^*$ and $w_{i,j} = 0$ for all i, j . This means that $\sum_{i=1}^{j^*} a_i(A) \geq \sum_{i=1}^{j^*} a_i(B)$. Repeating this process for all $j^* = 1, 2, \dots, n$, implies (2).

References

Baucells M, Sarin RK (2013) Guided decision processes. EURO J Decis Process 1:29–44

- Baucells M, Carrasco JA, Hogarth RM (2008) Cumulative dominance and heuristic performance in binary multi-attribute choice. *Oper Res* 56:1289–1304
- Brighton HJ (2006) Robust inference with simple cognitive models. In: Lebiere C, Wray R (eds) AAAI spring symposium: cognitive science meets AI hard problems, pp 17–22
- Carrasco JA, Baucells M (2008) Tight upper bounds for the expected loss of lexicographic heuristics in binary multiattribute choice. *Math Soc Sci* 55:156–189
- Chater N, Oaksford M, Nakisa R, Redington M (2003) Fast, frugal, and rational: how rational norms explain behavior. *Org Behav Human Dec Proc* 90:63–86
- Chipman H (1996) Bayesian variable selection with related predictors. *Canadian J Stat* 24:17–36
- Czerlinski J, Gigerenzer G, Goldstein DG (1999) How good are simple heuristics? In: Gigerenzer G, Todd PM, ABC research group (eds) Simple heuristics that make us smart. Oxford University Press, New York, pp 97–118
- Dawes RM, Corrigan B (1974) Linear models in decision making. *Psychol Bull* 81:95–106
- Gigerenzer G, Todd PM, ABC research group (1999) Simple heuristics that make us smart. Oxford University Press, New York
- Gigerenzer G, Hertwig R, Pachur T (2011) Heuristics: the foundations of adaptive behavior. Oxford University Press, New York
- Hogarth RM, Karelaia N (2005a) Ignoring information in binary choice with continuous variables: when is less “more”? *J Math Psychol* 49:115–124
- Hogarth RM, Karelaia N (2005b) Simple models for multiattribute choice with many alternatives: when it does and does not pay to face trade-offs with binary attributes. *Manag Sci* 51:1860–1872
- Hogarth RM, Karelaia N (2006a) Regions of rationality: maps for bounded agents. *Decision Anal* 3:124–144
- Hogarth RM, Karelaia N (2006b) “Take-the-best” and other simple strategies: why and when they work “well” with binary cues. *Theory Decis* 61:205–249
- Katsikopoulos KV (2011) Psychological heuristics for making inferences: definition, performance, and the emerging theory and practice. *Decision Anal* 8:10–29
- Katsikopoulos KV (2013) Why do simple heuristics perform well in choices with binary attributes? *Decision Anal* 10:327–340
- Katsikopoulos KV, Fasolo B (2006) New tools for decision analysts. *IEEE Trans Syst Man Cybern Part A Syst Hum* 36:960–967
- Katsikopoulos KV, Schooler LJ, Hertwig R (2010) The robust beauty of ordinary information. *Psychol Rev* 117:1259–1266
- Keeney RL, Raiffa H (1976) Decision-making with multiple objectives: preferences and value tradeoffs. Wiley, New York
- Kendall M (1938) A new measure of rank correlation. *Biometrics* 30:81–89
- Kirkwood CW, Sarin RK (1985) Ranking with partial information: a method and application. *Oper Res* 33:193–204
- Knight FH (1921) Risk, uncertainty and profit. Houghton Mifflin, Boston
- Li X, Sudarsanam N, Frey DD (2006) Regularities in data from factorial experiments. *Complexity* 11:32–45
- Martignon L, Hoffrage U (2002) Fast, frugal, and fit: simple heuristics for paired comparison. *Theory Decis* 52:29–71
- Martignon L, Katsikopoulos KV, Woike JK (2008) Categorization with limited resources: a family of simple heuristics. *J Math Psychol* 52:352–361
- Myers RH (1986) Classical and modern regression with applications. Duxbury Press, Boston
- Payne JW, Bettman JR, Johnson EJ (1993) The adaptive decision maker. Cambridge University Press, Cambridge
- Şimşek Ö (2013) Linear decision rule as aspiration for simple decision heuristics. In *Adv Neural Info Proc Syst* 26:2904–2912
- Tversky A (1972) Elimination by aspects: a theory of choice. *Psychol Rev* 79:281–299
- Wu CFJ, Hamada M (2000) Experiments: planning, design, and parameter optimization. Wiley, New York