On the estimation of a satisficing model of choice using stochastic multicriteria acceptability analysis

Ian Durbach
Department of Statistical Sciences, University of Cape Town, South Africa

Keywords: satisficing, stochastic multicriteria acceptability analysis, goal programming, simulation

Abstract

This paper addresses the task of estimating the kind of aspirations that are most likely to have given rise to an observed partial rank ordering of alternatives. The proposed approach uses a Tchebycheff goal programming formulation to serve as a descriptive model of choice and stochastic multicriteria acceptability analysis to generate candidate aspiration vectors and estimate the true aspiration vector from this candidate set. A simulation experiment is used to assess the accuracy of the estimated aspiration levels in a variety of problem contexts. The approach can perform well if strong performance is demanded on a small subset of the attributes and the rank order that is observed is sufficiently detailed.

1 Introduction

Descriptions of the way that trade-offs between conflicting criteria should and do occur have been the source of much conflict in the decision sciences. Two dominant frameworks exist to describe the way in which people make decisions when multiple objectives are involved; subjective expected utility (SEU, e.g. [35, 16]) and what might be called 'heuristic' methods (e.g. [24, 9]). In the SEU approach, a decision maker is assumed to possess a cardinal utility function which expresses his or her preferences over the set of options being considered, so that a rational decision is equated with selecting that alternative that provides maximum utility. In its simplest form, the utility U_i of alternative i to a particular individual is a latent or unobserved variable that is expressed as an additive function of a number of exogenous variables x_{ij}, so that

$$U_i = \sum_{j=1}^{J} \beta_j u_j(x_{ij}) + \epsilon_i \tag{1}$$

where $u_j(\cdot)$ is a function mapping attribute evaluations on attribute j to utilities. Although this paper is not concerned with the SEU model of decision-making, it is useful to make two observations that are supported by so much evidence that they approach statements of fact. Firstly, the axioms underlying SEU are known to be systematically and consciously violated by a wide variety of decision makers in a wide variety of problem contexts, the most famous examples of which are the so-called 'paradoxes' of Allais [1] and Ellsberg [7]. These violations require extensions to the basic SEU framework, for example prospect theory [14]. Secondly, the SEU model continues to be widely used in descriptive models of choice (e.g. in economics, [27, 13, 3]; and in marketing, [26, 28, 31, 4]).

An alternate framework for examining how people make decisions is provided by heuristic methods that stress the importance of being able to find alternatives that are 'good enough' (rather than ones maximising utility) because of naturally-occurring constraints on time, inclination, cognitive abilities, and other environmental variables that vary from problem to problem. The particular heuristic that is the focus of this paper is the well-known 'satisficing' heuristic [29] proposed by Herbert Simon as part of his bounded rationality program. In terms of the satisficing model, decision makers select an acceptable alternative by setting goals on the various aspects of the problem to be considered, and selecting the first alternative that satisfies the set of goals. This very static notion of decision making is made dynamic by allowing goals to be revised during the course of the search to be more demanding or more accommodating depending on the ease with which a satisficing alternative is found. The two key concepts underlying the satisficing model are therefore the setting of goals or aspirations and sequential search with possible revision of aspirations.

A satisficing approach to decision making has an intuitive descriptive appeal and has proved descriptively useful in a number of instances (see [30], chapter 2, or [8] for a more recent example) but has proved difficult to operationalise and implement with the same rigour that the SEU models achieve. There is a large body of literature on so-called ideal point methods (e.g. [32, 15, 21]) that evaluate by a weighted sum of deviations from an ideal point or vector of aspiration levels, but this is only indirectly relevant in that aspiration levels are also used. The underlying decision making process – a weighted additive model and the use of squared deviations both above and below a desired intermediate attribute value – is quite different from a model implied by satisficing. Possibly the best-known operationalisation of satisficing exists outside the sphere of descriptive decision making in the normative field of goal programming (for a recent application see [23]). Goal programming has provided an operational definition of a decision maker who sets goals or aspiration levels for each objective under consideration and then evaluates a prospective alternative by comparing the performance of that alternative to the set of aspirations using some distance measure. The alternative that is chosen minimizes the underachievement of the performance levels from the aspirations. The link between satisficing and goal programming was first explicitly identified by James Ignizio in [12] and has since been quite widely referenced by other writers (e.g. [10, 25, 2]).

In this paper a particular interpretation of satisficing is borrowed from goal programming and used as a descriptive choice model. Essentially, the paper addresses the task of discovering the kind of aspirations that are most likely to have given rise to an observed choice of alternative (or more generally, to a particular partial rank ordering of alternatives). The aspiration levels are estimated by Monte Carlo simulation using stochastic multicriteria acceptability analysis using achievement or reference functions (SMAA-A or Ref-SMAA [19, 5]). Traditionally, the SMAA family of methods has been used in a normative capacity to help provide information to a group of decision makers about the types of preference information that would lead to the selection of a particular alternative. This is done in such a way that the involvement of the decision makers in directly expressing their preference information can be located anywhere along a continuum from no direct assessment at all (although there is only one real-world application of this type [11]) to complete assessment of preferences as for regular normative decision aid (see [17, 20, 33]). In this paper, the low preference information nature of the SMAA-A method is exploited in order to apply it in a descriptive setting to estimate aspiration levels. The resulting estimate of the underlying aspiration vector could be used in a similar way to the output of existing descriptive choice models i.e. as a means of informing the strategies of other decision makers in a possibly game-theoretic environment.

The effectiveness of the proposed method is tested using a simulation experiment in which model accuracy is evaluated under a variety of simulated problem contexts. Monte Carlo simulation is thus used for two different purposes in this paper. Firstly, it is used by the SMAA method to generate possible preferences (i.e. aspiration vectors) that may have given rise to the rank order that is observed (as done in any SMAA application e.g. [18]). Secondly, it is used to generate hypothetical problem contexts in order to test the potential effectiveness of the approach and to identify any influences that it may be particularly sensitive to (e.g. [6]).

The remainder of the paper is structured as follows. After introducing some basic notation in section 2, section 3 describes a means of inferring the aspiration levels in a satisficing model of decision making. Section 4 describes a simulation experiment designed to investigate the effectiveness of this approach under various environmental conditions. In particular, the simulation experiment investigates the effect of different model applications (proportion of observed rank order used, errors in the assessment of the rank order) under different decision maker psychologies (the types of aspirations held) and different problem contexts (the number of alternatives and attributes). The results of the simulation experiment are presented and discussed in section 5, following which section 6 provides some concluding comments and directions for future research.

2 Notation

We refer to a set of I alternatives indexed by $i \in \{1, \dots, I\}$, and a set of J attributes indexed by $j \in \{1, \dots, J\}$ which are used as bases for evaluating the alternatives, so that by x_{ij} we denote the evaluation of alternative i on attribute j. Attributes have been defined without loss of generality so that more is always preferred. The aspiration level relating to the desired level of performance on attribute j is given by g_j, which can be assembled into a vector of aspirations $\mathbf{g} = [g_1, \dots, g_J]$. The set of all possible aspiration vectors is given by $\mathbf{G}$, which then represents every possible combination of goals that could be specified by the decision maker. Given the set of attribute evaluations x_{ij}

and aspirations $\mathbf{g}$, it is possible to compute the deviation δ_{ij} of each evaluation from the corresponding goal on that attribute i.e. $\delta_{ij} = g_j - x_{ij}$ for each attribute j. In cases where the deviational variables are negative, $-\delta_{ij}$ becomes a measure of overachievement.

3 Implementing satisficing as a model of choice

Alternatives are rank ordered according to a slightly adapted version of the Tchebycheff goal programming formulation in which alternatives are ranked from best to worst according to the minimization of

$$\Delta_i = \max_{j=1}^{J} w_j \delta_{ij} + \epsilon \sum_{j=1}^{J} w_j \delta_{ij} \qquad (2)$$

where w_j is a weight indicating the relative importance (to be interpreted in a swing-weight sense) of attribute j, and ϵ is a small positive constant that ensures that the Pareto-optimal alternative is selected in cases where two alternatives have the same maximum weighted deviation. This particular form of the Tchebycheff formulation is given in [2], where it is identified as closely aligned with satisficing. In a continuous setting the above formulation implies that the decision maker would attempt to improve performance on that attribute j^- that is currently worst-performing i.e. has the largest value of $w_j \delta_{ij}$, and will not consider performance on the other attributes until j^- is no longer the worst-performing attribute. At this stage, attention would shift to the new most-unsatisfactory attribute. An important difference distinguishing this approach from a traditional lexicographic strategy is that in the former attention may shift while a goal is still unfulfilled – though the approach can be described in terms of the more general extended lexicographic goal program [25]. In a discrete setting, the implication is that in cases in which not all aspirations are satisfied, an alternative will be evaluated by its performance on the attribute on which underachievement is greatest. If all aspirations are satisfied, the alternative is evaluated by its performance on the attribute on which overachievement is smallest. This particular interpretation of satisficing is superficially quite different from the original notion of satisficing described in the introduction, but can be reconciled by taking a view that places emphasis on the dynamic nature of goal-setting i.e. one that supposes that in cases where aspirations are easily fulfilled, a natural response is to increase aspirations and re-evaluate alternatives.

Given a particular observed rank order (which may not necessarily be complete) and the model of satisficing described above, we wish to estimate the aspiration levels that might have given rise to the rank order. This estimation can be carried out using Monte Carlo simulation and an application of SMAA-A. The first step in inferring aspiration levels is the identification of the set of aspiration vectors $\mathbf{G}_\Gamma$ that are compatible with the observed preference order Γ. Since aspirations may be specified to an arbitrary precision, all non-dominated alternatives possess an infinite number of aspiration vectors that would make that alternative the preferred one, and in practice the infinite set $\mathbf{G}_\Gamma$ is represented by a representative set of compatible aspiration vectors $\hat{\mathbf{G}}_\Gamma$. This set of compatible aspiration vectors $\hat{\mathbf{G}}_\Gamma$ is constructed by a Monte Carlo simulation which at each iteration randomly generates an aspiration vector $\mathbf{g}$ and applies equation (2) to arrive at a complete predicted rank order. If this predicted rank order is consistent with the (possibly incomplete) observed rank order, then $\mathbf{g}$ is included in $\hat{\mathbf{G}}_\Gamma$. If not, it is discarded. The exact number of Monte Carlo iterations that are required to achieve a given level of estimation accuracy has been shown in [34] to depend on the actual quantity that is being estimated. In order to estimate the expected aspiration vector (or central weight vector in conventional SMAA) within ξ of the true aspiration vector with 95% confidence, one requires $1.96^2/4\alpha\xi^2$ iterations, where α is the proportion of iterations contributing to the set of consistent aspiration vectors i.e. the proportion of generated aspiration vectors that are not discarded. For example, in order to achieve error limits of 0.05 on the estimated aspirations under the assumption that the output from one in every 250 iterations is retained, 96040 iterations are required.

The set of compatible aspiration vectors $\hat{\mathbf{G}}_\Gamma$ provides all possible combinations of aspiration levels that might have given rise to the observed rank ordering of alternatives. Given the current state of knowledge, the satisficing decision maker is using *one* of these aspiration vectors – but it is not clear which one. Intuitively, what is desired is some kind of single best estimate of the currently-held aspirations, as well as an indication of how variable or uncertain this estimate is. Some guidance on this problem is provided by SMAA-A, which has grappled with a similar problem of providing summary measures of $\hat{\mathbf{G}}_\Gamma$ to decision makers that they will find useful in moving toward a final decision.

First, a natural estimate for the set of aspirations currently held by the decision maker is the centre of gravity of the hypervolume of compatible aspiration vectors, the so-called *central aspiration vector* $\mathbf{g}^c$ [17]

$$\mathbf{g}^c = \int_{\mathbf{G}_\Gamma} \mathbf{g} \, dg \tag{3}$$

The central aspiration vector defines a single summarised average of all compatible aspiration vectors, and represents the expected preferences of a typical decision maker exhibiting the observed preference order. In practice, the central aspiration vector is simply computed as the centre of gravity of all those vectors in $\hat{\mathbf{G}}_\Gamma$.

Second, an indication of the variability around the central aspiration vector summary measure can be provided by two of the other summary measures popular in SMAA – the acceptability index and the minimum and maximum compatible aspiration vectors [17]. The acceptability index is an indication of the relative size of the hypervolume representing the collective set of aspiration vectors that are compatible with the observed rank order Γ,

$$A(\Gamma) = \frac{\text{vol}(\mathbf{G}_\Gamma)}{\text{vol}(\mathbf{G})} = \frac{\int_{\mathbf{G}_\Gamma} dg}{\int_{\mathbf{G}} dg} \tag{4}$$

The acceptability index has a minimum value of 0 (where any dominated alternative appears above its dominating alternative in the rank order) and a maximum of 1 (where every alternative is dominated by all the alternatives that lie above it in the observed rank order). Since in this latter case the ranking of alternatives will be Γ regardless of what aspiration levels are used, there is total uncertainty about the central aspiration vector. As $A(\Gamma)$ decreases from 1 to 0, the volume containing the set of compatible aspiration vectors decreases in size and thus, since aspiration vectors are randomly generated by the Monte Carlo simulation, certainty about the central aspiration vector increases. Uncertainty is thus an increasing function of the acceptability index $A(\Gamma)$ in the region $[0, 1]$, and is undefined elsewhere.

While the acceptability index provides a useful summary of overall uncertainty, it may also be useful to evaluate the range of aspirations on each attribute that may give rise to the observed rank order. The minimum and maximum compatible aspiration levels for each attribute j are defined by

$$\mathbf{g}_j^{min} = \min_{\mathbf{g} \in \mathbf{G}_\Gamma} g_j \quad (5)$$

$$\mathbf{g}_j^{max} = \max_{\mathbf{g} \in \mathbf{G}_\Gamma} g_j \quad (6)$$

and can be used to obtain a bound on the aspirations that still result in a rank order that is compatible with what was observed. Again, in practice the ranges of compatible aspiration vectors are simply computed from the vectors in $\hat{\mathbf{G}}_\Gamma$.

The approach described above can be summarised as an application of two operational research techniques usually reserved for normative decision making, Tchebycheff goal programming and SMAA-A, to estimate aspiration levels in a satisficing model. Aspiration vectors are generated at random in a Monte Carlo simulation and the rank order generated from using the aspirations in a Tchebycheff goal program is evaluated against the observed rank order. If they are consistent with one another, the aspiration vector is retained in the set of 'possibly true' aspirations. Finally, summary measures from SMAA-A are applied to the set of compatible aspiration vectors to provide a single estimate of the aspiration levels and two measures which indicate the extent to which that estimates is uncertain. Given the heuristic nature of the approach and estimation, it is important to provide some indication about how accurate the aspiration levels estimated from the SMAA-A procedure outlined above might be at modelling the true preferences of a hypothetical decision maker. In the following section, an attempt is made to answer this question using a simulation experiment.

4 A simulation-based test of the approach

At first glance it may appear strange to test a descriptive model of choice using a simulation experiment. While a real-world study of the proposed approach is a hope for future research, a simulation experiment provides a useful starting point to such a study and allows one to evaluate whether in principle the proposed approach might accurately represent the true underlying aspiration levels of those who are using a satisficing approach to decision making. The general simulation structure that is used comprises four elements:

1. Creating a problem context comprising the attribute evaluations and weights.
2. Generating the idealised and observed rank orders of alternatives based on an underlying aspiration vector that is possessed by the decision maker.
3. Applying the proposed approach to generate a central aspiration vector to be used as an estimate of the underlying aspirations.
4. Comparing the central aspiration vector to the true aspiration levels to evaluate the accuracy of the proposed approach.

4.1 Creating the problem context

The simulated decision problem consists of I alternatives evaluated over J attributes, with the evaluations of the performance level of alternative i on attribute j, x_{ij}, generated to lie on the unit hypersphere in order to guarantee that all alternatives are non-dominated. Without loss of generality, attribute evaluations are defined in a maximising sense i.e. more is better, and standardised to have a maximum attribute value of 1 and a minimum attribute value of 0. Since attribute weights in a goal programming context are predominantly used to ensure attribute values are scaled appropriately, and all attributes are already simulated to lie between 0 and 1, equal attribute weights are used throughout.

4.2 Generating the idealised and observed rank orders of alternatives

The simulated decision maker is assumed to use a satisficing strategy in the sense implied by equation (2). That is, he or she possesses a 'true' underlying preference structure which can be represented by a vector of aspiration levels $\mathbf{g}^T$, and alternatives are evaluated based on the extent to which their performance differs from this set of aspiration levels. In the simulations, five different types of decision maker are modelled by varying the nature of the aspiration levels. Table 1 shows the five different types of aspirations that are used. They are intended to set roughly more challenging estimation tasks for the proposed approach, and make a distinction between attributes for which aspiration levels are 'high' and those for which aspiration levels are 'low'. This allows qualitatively different decision makers to be simulated. For Type 1 aspirations, aspiration levels are 'high' on only one attribute and 'low' on all others, and the gap between 'high' and 'low' is substantial, representing a decision maker who cares only about performance on one attribute to the near-exclusion of all others. For Type 2 aspirations not only are two aspiration levels 'high', but the gap between 'high' and 'low' is also narrower. Type 3 aspirations again both increase the number of high aspiration levels while narrowing the gap between aspirations on all attributes. Type 4 and Type 5 aspirations probe whether the approach is firstly able to detect true compromise decision making, and secondly able to detect the difference between a demanding and easily-satisfied decision maker (Types 4 and 5 respectively). Let the set of attributes with 'high' aspiration levels be denoted by H and the set of attributes with 'low' attributes be denoted by L.

Aspiration	Number of 'high' g_j	Aspirations generated from	Number of 'low' g_j	Aspirations generated from
Type 1	1	$U[0.8, 1]$	$J-1$	$U[0.1, 0.3]$
Type 2	2	$U[0.7, 0.9]$	$J-2$	$U[0.2, 0.4]$
Type 3	4	$U[0.6, 0.8]$	$J-4$	$U[0.3, 0.5]$
Type 4	J	$U[0.35, 0.55]$	–	–
Type 5	–	–	J	$U[0.15, 0.25]$

Table 1: Aspirations types used in the simulation

Although the aspiration levels remain unobserved and are perhaps even quantities that the decision maker is not consciously aware of, they are assumed to represent his or her true underlying preferences, and give rise to a rank ordering of alternatives from best to worst. This is termed the idealised or 'true' rank order.

While the idealised rank order is by definition error-free, the fact that a decision maker may not be perfectly aware of his or her underlying preferences means that it is possible that the rank order that is specified by the decision maker might not be the same as the 'true' rank order implied by their 'true' underlying preferences. In particular, it seems reasonable to suggest that decision makers become less certain of their preferences the further down the rank order one goes, so that they are likely to be more certain of their favourite alternative than their 8th or 9th favourite alternative. This conjecture is simulated by considering a probability of rank reversal for each rank r, indicating the probability that the alternative in position r in the current rank order is mistakenly swapped with the alternative in position $r+1$ to form a new rank order. This process is performed iteratively starting at $r = 1$ and ending at $r = I - 1$, so that it is possible for a alternative to move more than a single rank. The previous discussion suggests that the probability of rank reversal should increase from best- to worst-position in the observed rank order, so that the probability is modelled simply by $r\omega$, where ω is a parameter of the simulation that can take on the values 0 (a "no-error" condition) or 0.08 (an "error" condition). The error condition parameter was chosen on the basis of trial-and-error experimention to provide a fairly high overall propensity to make mistakes (on average 2.8 reversals when $n = 8$ and 10.2 reversals when $n = 16$) while offering relatively good assessment of the most-preferred ranks (the top two alternatives are correctly assessed in 77% of all cases).

Finally, it may not always be possible to base the comparison of the simulated and observed rank orders on all I positions in the rank order. Firstly, it may be that only an incomplete rank ordering of alternatives is observed (in an extreme but potentially quite common case, just the selected alternative might be observed). Furthermore, both increases in computational effort and the possible errors in the observed rank order suggest that it might not necessarily be beneficial in any case to base the comparison on all positions, but only on the first ϕI positions in the rank order, where ϕ is a parameter to be investigated by the simulation. The resulting rank order is termed the 'observed' rank order and it is only this rank order that is available to the satisficing model.

4.3 Applying the satisficing model

The simulation of the satisficing model begins by randomly generating an aspiration vector $\mathbf{g}$ and solving the resulting decision problem using (2) to arrive at a simulated rank order. If the simulated rank order is consistent with the observed rank order, then that aspiration vector is added to the set of compatible aspiration vectors $\hat{\mathbf{G}}_\Gamma$. A total of 100 000 iterations are used in each condition, which is sufficient to give an accuracy of 0.05 on the estimated aspirations under the assumption that the output from one in every 250 iterations is retained [34]. Initial experiments indicated minimal change in results beyond this number of iterations. After 100 000 iterations the central aspiration vector is computed for the set of compatible aspiration vectors and used as an approximation of the decision maker's true aspiration levels.

4.4 Evaluating the results of the satisficing model

The main output of the previous step is a central aspiration vector which represents an estimate of the true aspiration vector possessed by the decision maker. The evaluation of how well the satisficing model performs is based on a comparison of these two vectors in the form of the following four measures:

- d_{avg}: the average absolute difference over all attributes between the true aspiration vector and the central aspiration vector i.e. $d_{avg} = 1/J \sum_{j=1}^{J} |g_j^T - g_j^c|$.
- d_{max}: the maximum absolute difference over all attributes between the true aspiration vector and the central aspiration vector i.e. $d_{max} = \max_j |g_j^T - g_j^c|$
- V_1: an indicator variable taking on a value of 1 if the estimated aspiration levels for all those attributes with *a priori* 'high' aspiration levels exceed the estimated aspiration levels for all those with *a priori* 'low' aspiration levels i.e. if $g_j^c > g_k^c, \forall j \in H, k \in L$, and 0 otherwise. Note that this measure is not used for aspiration types 4 and 5.
- V_2: an indicator variable taking on a value of 1 if $V_1 = 1$ and, in addition, the estimated aspiration levels of all 'high' attributes are above the lower bound of the

interval used to generate 'high' aspiration levels of that type, g^A i.e. if $g_j^c > g^A$ and $g_j^c > g_k^c, \forall j \in H, k \in L$, and 0 otherwise. Note that this measure is also not used for aspiration types 4 and 5.

- S: an indicator variable taking on a value of 0 if the set of favourable aspiration vectors is still empty after 100 000 iterations i.e. no central aspiration vector can be computed, and 1 otherwise.

A summary of the effects investigated and the parameter values simulated is shown in Table 2. Numbers of attributes ($m = 3, 5, 9$) were chosen in line with previous research [5] as well as what seemed to represent plausible limits for descriptive choice (e.g. [22]). Numbers of alternatives ($n = 8, 16$) were chosen so as to allow for meaningful differences in the proportions of the full rank order considered to be investigated. For each combination of parameters, 100 simulations were run. This was sufficient for the standard errors of the mean d_{avg} values to generally lie between 0.002 and 0.005 and the standard errors of the mean d_{max} values to generally lie between 0.005 and 0.01. Although the results presented below do not make much reference to statistical significance, Shapiro-Wilk's tests showed that the distributions of both $\bar{d}_{avg}$ and $\bar{d}_{max}$ are at least approximately normal in the vast majority of simulated problem contexts. As a result, any difference in $\bar{d}_{avg}$ in excess of 0.02 and any difference in $\bar{d}_{max}$ in excess of 0.04 can be considered to be significant at the 5% level at least.

Parameter	Values	Description
I	$8, 16$	Number of alternatives
J	$3, 5, 9$	Number of attributes
Aspirations	Type 1, 2, 3, 4, 5	Defines the aspiration levels
ω	$0, 0.08$	Induces errors in the observed rank order
ϕ	$1/I, 1/4, 1/2$	Proportion of rank order considered

Table 2: Parameter values for the simulation experiment

Before turning to the results themselves, it may be beneficial to sketch some hypothesised relationships. Other things being equal, the size of the hypervolume of favourable aspiration vectors for a particular alternative is decreased by increasing the number of alternatives, decreasing the number of attributes, and by increasing the proportion of the rank order considered. These effects should therefore increase the accuracy of estimation. Furthermore, it seems clear that assessment errors will negatively impact on model accuracy, particularly when a large proportion of the rank order is used. No clear hypotheses are suggested for the effect of aspiration type.

5 Simulation results

The analysis of results is divided into three sections: the effects of aspiration type and the proportion of the rank order considered; the size of the decision problem; and any errors in the assessment of the observed rank order of alternatives. The results of a multivariate analysis of variance are provided in Table 3 as a quick and concise way to summarise the statistical significance of the main effects and all second- and third-order interactions. The MANOVA excludes the outcome variables V_1 and V_2, since these are not calculated for all aspiration types. Type 3 aspirations are also excluded since these are not possible when $m = 3$. Significance at the 5% and 0.5% level is denoted by a single and double asterisk respectively.

All main effects are highly significant and most of the interaction effects of second- and third-order are also significant at the 0.5% level. Our focus will mainly be on those interactions that are most significant; specifically on the two-way interactions between aspiration type and each of the number of alternatives n, number of attributes m, and proportion of rank order considered ϕ, as well as on the three-way interaction between

Effect	DoF	F	Effect	DoF	F
n	2	135.2**	Asp$\times\phi$	12	151.3**
m	4	1443.9**	$\omega \times \phi$	4	43.5**
Asp	6	1518.5**	$n \times m\times$Asp	12	32.7**
ω	2	132.7**	$n \times m \times \omega$	4	0.8
ϕ	4	180.8**	$n\times$Asp$\times\omega$	6	1.5
$n \times m$	4	29.7**	$m\times$Asp$\times\omega$	12	1.8*
$n\times$Asp	6	139.4**	$n \times m \times \phi$	8	12.8**
$m\times$Asp	12	382.7**	$n\times$Asp$\times\phi$	12	19.2**
$n \times \omega$	2	24.8**	$m\times$Asp$\times\phi$	24	63.0**
$m \times \omega$	4	3.0*	$n \times \omega \times \phi$	4	7.8**
Asp$\times\omega$	6	11.6**	$m \times \omega \times \phi$	8	1.9*
$n \times \phi$	4	82.5**	Asp$\times\omega \times \phi$	12	5.8**
$m \times \phi$	8	74.1**			

Table 3: Multivariate ANOVA results for the simulation experiment

aspiration type, m and ϕ.

5.1 Effect of aspiration type and proportion of rank order considered

We begin by considering the task of inferring satisficing levels in a baseline case of 5 attributes and 8 alternatives where there are no errors in the observed preference ordering of alternatives. Results are shown in Table 4, which shows the effect of considering different proportions of the rank order when constructing the set of favourable aspiration vectors. Figure 1 shows the proportion of simulations in which (a) the estimated aspiration levels for all *a priori* 'high' attributes are above the estimated aspiration levels for all *a priori* 'low' attributes, and (b) the estimated aspiration levels for all 'high' attributes are above the estimated aspiration levels for all 'low' attributes *and* are sufficiently high to be clearly identified as such. Note that since no high/low distinction is made for aspiration types 4 and 5, these do not form part of Figure 1. Also, in all cases the set of compatible aspiration vectors was non-empty following 100 000 iterations i.e. a solution was found, and so these results are not shown in Table 4.

The main finding emerging from the results in Table 4 and Figure 1 can be summarised

Aspir.	ϕ	$\bar{d}_{avg}$ 1/I	0.25	0.5	$\bar{d}_{max}$ 1/I	0.25	0.5
Type 1		0.21	0.13	0.09	0.39	0.26	0.19
Type 2		0.15	0.11	0.08	0.28	0.24	0.17
Type 3		0.16	0.15	0.12	0.29	0.28	0.27
Type 4		0.11	0.12	0.14	0.22	0.26	0.27
Type 5		0.31	0.32	0.34	0.47	0.50	0.51

Table 4: Effect of proportion of rank order observed, ϕ (when $n = 8, m = 5$, and $\omega = 0$)

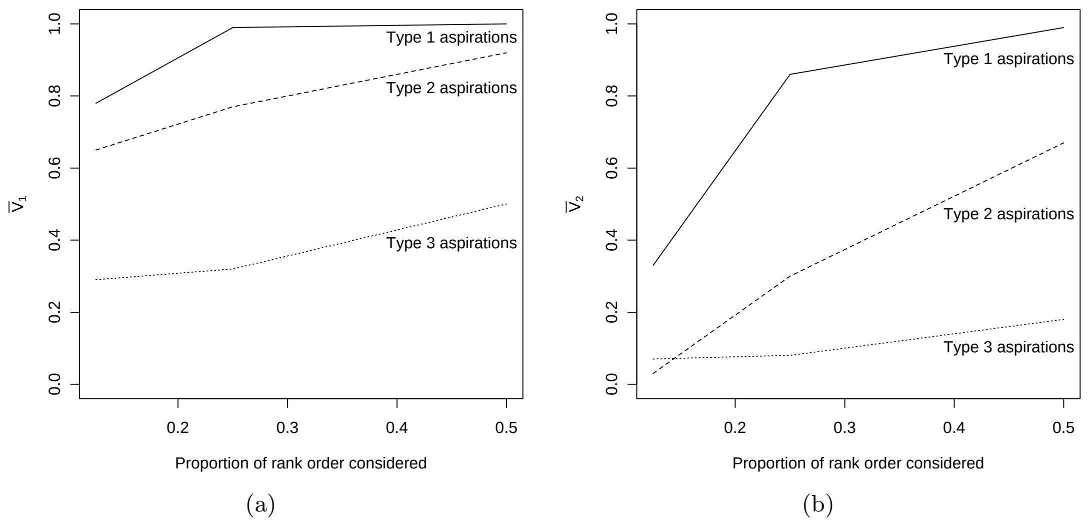


Figure 1: Effect of proportion of rank order considered on (a) V_1: the ability to adequately discriminate 'low' from 'high' aspiration levels; (b) V_2: the ability to discriminate and also to identify all 'high' aspiration levels as 'high'

as follows:

Result 1: *True satisficing levels can be well approximated provided that two conditions are met: (1) there are one or two attributes for which strongly better performance is demanded, (2) a sufficiently detailed rank order is used. If either of these conditions are not met, accuracy is poor*

The MANOVA results in Table 3 indicate a strong interaction effect between the proportion of the rank order used and aspiration type ($F = 151.3, p < 0.005$). Figure 1 shows that in the case of there being only one attribute with a clearly higher aspiration level (Type 1 aspirations), an approach considering only the first 25% of the rank order is able to estimate the aspiration level for the single 'high' attribute as lying above all other attributes in 99% of simulations. In addition it is also able to identify the aspiration level as suitably 'high' (above 0.8) in 86% of simulations. For cases in which the decision maker has two attributes where clearly better performance is desired (Type 2 aspirations), it is necessary to use half of the full rank order ($\phi = 0.5$) in order to obtain good approximation accuracy. If this is done, the estimated aspiration levels on both 'high' attributes lie above all other attributes in 92% of simulations, and are identified as suitably 'high' (above 0.7) in 67% of simulations. Table 4 supports these results, showing that for Type 1 and 2 aspirations, average deviations below 0.1 and maximum deviations of less than 0.2 can be obtained without knowledge of the full preference order.

In contrast, the approach is less successful in estimating aspiration levels for the other three aspiration types, in which more compromising solutions are sought. The poor performance for Type 3 aspirations is due to the difficulty of the estimation task, which requires that the approach detect the small discrepancy between the upper bound of the interval used to generate 'low' aspiration levels (0.5) and the lower bound of the interval used to generate 'high' aspiration levels (0.6). Given the approximation errors in Table 4 and the small difference between 'high' and 'low' aspirations, it is likely that in a large proportion of simulations at least some confusion of attributes occurs, though values of $\bar{d}_{avg}$ appear to remain reasonable provided that half of the rank order is considered. In the case of the true 'compromise' aspirations represented by Type 4 and 5, the approach appears largely unable to detect that a similar level of performance is desired on all attributes. This is reflected in Table 5, which shows that the difference between the highest and lowest aspiration levels is systematically overestimated by the approach for both aspiration types 4 and 5. This overestimation increases as the true range becomes narrower. In particular, the approach does not clearly distinguish between Type 3 aspirations, where the average range of aspirations is 0.36, and Type 4 aspirations, where the average range is more than 60% smaller at 0.13. In addition, it would not in general be possible to infer whether the compromising decision maker was easily-satisfied (Type 5) or not (Type 4). As will be clear from the next section though, this is highly dependent on the number of alternatives and attributes employed.

Aspirations	True	$\phi = 1/I$	$\phi = 0.25$	$\phi = 0.5$
Type 1	0.76	0.48	0.62	0.67
Type 2	0.58	0.42	0.51	0.56
Type 3	0.36	0.35	0.41	0.42
Type 4	0.13	0.30	0.36	0.38
Type 5	0.06	0.30	0.33	0.34

Table 5: Mean difference between highest and lowest aspiration level

5.2 Effect of numbers of alternatives and attributes

The net change in the outcome measures that arises when increasing the number of alternatives from the baseline case discussed above (in which $n = 8$) to $n = 16$ and the number of attributes (from $m = 5$) to $m = 3$ or $m = 9$ are shown in Tables 6 and 7 respectively, together with the proportion of simulations in which solutions could be found.

Again, it is useful to summarise the results above in the following two findings before turning to a more detailed discussion:

Result 2: *Increasing the size of the alternative set significantly improves estimation*

		$\bar{d}_{avg}$			$\bar{d}_{max}$			$\bar{S}$	
Aspir.	ϕ 1/I	0.25	0.5	1/I	0.25	0.5		0.25	0.5
Type 1	–0.04	–0.09	–0.05	–0.01	–0.15	–0.08		1	1
Type 2	–0.02	–0.05	–0.03	+0.00	–0.10	–0.07		1	1
Type 3	+0.00	–0.03	+0.01	+0.00	–0.04	–0.01		1	0.75
Type 4	+0.01	+0.04	+0.05	+0.02	+0.02	–0.05		0.98	0.57
Type 5	+0.00	+0.00	–0.07	+0.03	–0.02	–0.20		0.99	0.36

Table 6: Mean changes in $\bar{d}_{avg}$ and $\bar{d}_{max}$ and proportion of solvable simulations when $n = 16$ alternatives are used ($m = 5$)

		$m = 3$			$m = 9$		$\bar{S}$	
Aspir. ϕ	1/I	0.25	0.5	1/I	0.25	0.5	0.25	0.5
Type 1	–0.08	–0.04	–0.02	+0.06	+0.09	+0.06	1	1
Type 2	+0.00	+0.00	+0.00	+0.03	+0.02	+0.03	1	1
Type 3	n/a	n/a	n/a	–0.04	–0.03	+0.00	1	1
Type 4	+0.01	+0.01	+0.00	–0.02	–0.02	–0.01	1	1
Type 5	–0.07	–0.15	–0.21	+0.00	+0.00	+0.00	1	1

Table 7: Mean changes in $\bar{d}_{avg}$ when $m = 3$ or $m = 9$ attributes are used ($n = 8$)

accuracy for Type 1, 2 and 5 aspirations but can cause difficulties in finding a solution for more comprimising aspiration types

Result 3: *Estimation accuracy decreases significantly as the number of attributes increases for Type 1, 2 and 5 aspirations. Where $m = 3$ attributes are employed, the approach shows a dramatic increase in its ability to identify comprimising preferences.*

The number of alternatives available to the decision maker has a fairly substantial effect on the accuracy of the central aspiration vector and as hypothesised the approach does better with the larger set of alternatives by on average reducing the area bounded by the favourable aspiration vectors. There is a strong interaction between the number of alternatives and the type of aspirations that are held ($F = 139.4, p < 0.005$), with the improvements in accuracy due to the larger alternative set being strongest under Type 1, 2 and 5 aspirations. Although estimation of Type 5 preferences improves substantially when $\phi = 0.5$, estimation accuracy remains poor with average deviations still above 0.25.

A potential problem with using a larger set of alternatives is the increase in the number of simulations in which no solution could be found because the set of favourable aspiration vectors was still empty after 100 000 iterations. This is particularly severe in the case of $\phi = 0.5$, where the proportion of unsolved simulations jumps to 25%, 43% and 64% for Type 3, 4 and 5 aspirations respectively. In contrast, the ability of the approach to find a solution in the case of Type 1 and 2 aspirations appears to be robust to changes in the size of the alternative set.

The hypothesised inverse relationship between model accuracy and the number of attributes used is also generally supported by the results shown in Table 7. However, there are some important interaction effects (between the number of attributes and aspiration type, $F = 382.7, p < 0.005$; and between attribute number, aspiration type and the proportion of the rank order considered, $F = 63.0, p < 0.005$) that have two important implications. Firstly, the inverse relationship between model accuracy and the number of attributes is strong for Type 1 and Type 5 aspirations and exists to a lesser extent for Type 2 aspirations, but is not observed for the other aspiration types. Importantly, the improvement in estimation accuracy under Type 5 aspirations when $m = 3$ is large enough that the approach is able to generate quite satisfactory results. When $\phi = 0.5$ average deviations drop to 0.09 and maximum deviations to 0.18, which is superior to performance with Type 3 and 4 aspirations and comparable to performance with Type 2 aspirations. The approach thus shows some ability to identify comprising preferences provided that the number of attributes is small. Secondly, for Type 5 aspirations, changes in the numbers of attributes used have far greater impact when ϕ is large. Changes in attribute type show little or no interaction between ϕ and the number of attributes used.

Although the average and maximum approximation errors in general increase significantly as the number of attributes increases, the approach continues to distinguish between those attributes with higher aspiration levels and those with lower aspiration levels. Provided that half the rank order is used, the estimated aspiration levels on 'high' attributes lie above all other attributes in 99% and 90% of simulations for Type 1 and 2 aspirations respectively, and are identified as 'high' (requiring an estimated aspiration level of above 0.8 for Type 1 and 0.7 for Type 2 aspirations) in 98% and 78% of simulations. This level of performance is similar to what is achieved when $m = 5$ attributes are used. Increasing the number of attributes has no effect on the ability of the model to find a solution.

5.3 Effect of errors in the observed preference order

Since the only preference information that is required of the decision maker is a partial rank ordering of the alternatives from best to worst, an interesting question with which to conclude this analysis is how sensitive the results are to errors in the assessment of this rank order. Such errors might creep into an experimental setting where the decision maker does not take sufficient care to fully represent his or her true preferences, or into an observational setting where small situational factors alter a decision maker's usual preference ordering. Table 8 shows the changes resulting from errors in the assessment of the observed rank order.

			$\bar{d}_{avg}$			$\bar{d}_{max}$		$\bar{S}$	
Aspir.	ϕ	$1/I$	0.25	0.5	$1/I$	0.25	0.5	0.25	0.5
Type 1		+0.00	+0.02	+0.06	+0.00	+0.05	+0.13	0.89	0.42
Type 2		+0.00	+0.02	+0.04	+0.00	+0.03	+0.10	0.93	0.54
Type 3		+0.00	+0.01	+0.01	–0.01	+0.02	+0.02	0.95	0.56
Type 4		+0.00	+0.00	–0.01	+0.01	+0.01	+0.03	0.90	0.52
Type 5		+0.00	+0.02	+0.07	+0.00	+0.05	+0.12	0.86	0.46

Table 8: Mean changes in accuracy and proportion of solvable simulations when errors are made

The main effect of errors in the assessment of the ranking of alternatives can be summarised as:

Result 4: *Errors in the assessment of the observed rank order operate by (a) increasing the probability that no solution can be found and (b) degrading the accuracy of the solution where one can be found.*

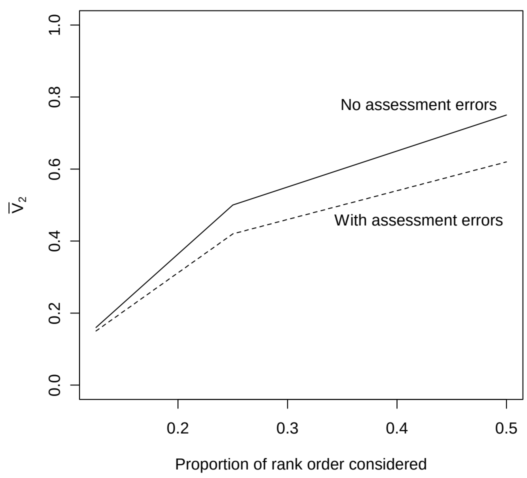
Result 5: *Deteriorations in estimation accuracy increase with the proportion of the rank order considered and the size of the alternative set, and can become substantial when both of these contributing factors take on their largest values*

The increased risk of not being able to find a solution is only pronounced when a large proportion of the rank order is being considered i.e. when $\phi = 0.5$. Increasing the number of simulations is unlikely to help in this case. Given the attribute evaluations, the reason for the empty set of compatible aspiration vector would most likely be that no aspiration vectors that are compatible with the erroneous rank order exist rather than the simulations not locating any within 100 000 iterations. This possibility is clearly exacerbated in cases where a higher value of ϕ implies a more detailed construction, since the set of favourable aspiration vectors will necessarily be smaller to begin with.

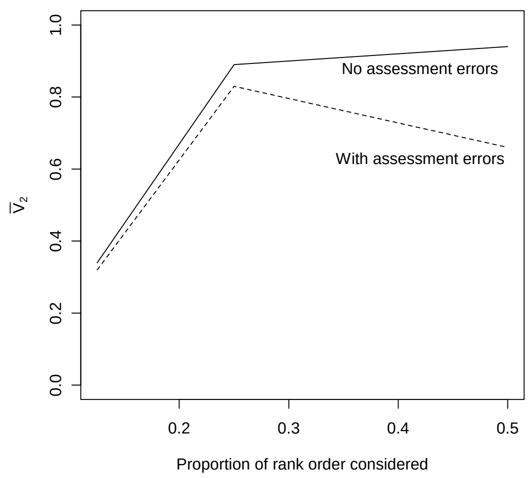
Where a solution can be found, the quality of this solution can differ quite widely. In most circumstances, the accuracy of the resulting central aspiration vector is only marginally worse than when no errors are made. Table 8 shows that deteriorations in d_{avg} and d_{max} are at most 0.02 and 0.05 respectively whenever $\phi < 0.5$. Further significant interactions exist between the effect of assessment errors and each of number of alternatives ($F = 24.8, p < 0.005$) and proportion of the rank order considered ($F = 43.5, p < 0.005$). Specifically, if there are a large number of alternatives ($n = 16$) and a large proportion of the rank order is considered ($\phi = 0.5$), assessment errors have far greater detrimental effects. Figure 2 considers only Type 1 and 2 aspirations, and shows the average accuracy statistic V_2 for both sizes of alternative sets as well as all proportions of the considered rank order. In cases where $n = 16$ and $\phi = 0.5$, the model correctly and unambiguously identifies all important attributes in only 66% of simulations in which any solution was identified. Note that this is nearly 20% *worse* that if a smaller proportion of the rank order is considered i.e. $\phi = 0.25$. The message that emerges is clear: in cases in which many alternatives are present, caution should be exercised in the use of a large proportion of the assessed rank order. If there is reason to believe that there are errors in the assessed rank order, then it is safer to make use of a smaller proportion.

6 Conclusions

The main contribution of this paper has been to propose and evaluate the possibility of using a combination of a Tchebycheff goal programming formulation and stochastic multicriteria acceptability analysis to implement a satisficing model of descriptive choice. The approach uses Monte Carlo simulation to randomly generate sets of aspirations and observes whether the rank orders obtained by substitution of the aspirations into a



(a) n=8



(b) n=16

Figure 2: The interaction effect between assessment errors and the proportion of the rank order considered on V_2 (a) when $n = 8$ alternatives are present; (b) when $n = 16$ alternatives are present. Both figures are for Type 1 and 2 aspirations only.

Tchebycheff goal program are consistent with an *a priori* observed rank order that is not necessarily complete. In such a way, a bank of 'possible' aspiration vectors are built up that are consistent with the observed choices and preferences. The SMAA-A technique is used to provide an estimate of the aspiration levels and two measures that indicate the extent to which the estimate is uncertain.

A simulation experiment was used to test the ability of the proposed approach to accurately estimate the central aspiration vector of a hypothetical decision maker (simulated in terms of the types of aspirations that might be held) under various problem contexts (number of alternatives and attributes present) and variations of the application of the approach (proportion of the observed rank order used and errors in the assessment of the rank order). The main findings emerging from the simulation experiment are:

1. The proposed approach appears to provide a good approximation in certain simpler

decision contexts, particularly when (a) the number of attributes is fairly small, (b) the decision maker demands good performance on a small subset of the attributes and accepts poor performance on the others, rather than searching for average performance across all attributes, and (c) the number of alternatives is of at least moderate size. Compromising preferences can be detected if a very limited number of attributes is used, but if more attributes are used the approximation results are not good. General use of the approach would seem best suited to the identification of the one or two attributes that are really driving the choice of a particular alternative. This does not preclude the approach from being potentially useful. The notion of bounded rationality suggests that in many circumstances, particularly in everyday decisions, decision makers will focus on just a small number of key attributes.

2. Approximation accuracy improves as a greater proportion of the observed rank order is utilised to construct the set of favourable aspiration vectors and, for those aspiration types for which the approach works well (see above), this improvement does not affect the probability of being unable to locate a solution. When aspirations on a single attribute dominate, estimation can be accurate even if only the top 25% of the rank order is examined. If conditions are less ideal, for example if aspirations on any other attributes are also high or if more attributes are used, then it is worth attempting to observe a greater proportion of the rank order.

3. The approach appears to be fairly robust to errors in the assessment of the observed rank order except in those cases where a substantial proportion (e.g. half) of a rank ordering of a large number of alternatives is being considered. In cases in

which both the number of alternatives and proportion of the rank order used is large, assessment errors can result in fairly severe decreases in model accuracy. In circumstances in which assessment errors might possibly occur, a sensible approach is to use a smaller proportion of the rank order, since in all such cases results remain good in spite of any errors in assessment. Apart from degrading model accuracy, assessment errors also increase the probability that a solution cannot be found at all, again particularly in cases where a substantial proportion of the observed rank order is used.

In assessing opportunities for future development, there seem to be two main directions to take. Firstly there might be substantial improvements in estimation accuracy that could be achieved by making the Monte Carlo simulation more efficient in generating compatible aspiration vectors; potentially by selecting initial aspiration vectors that are close to the performance of the *a priori* preferred alternative and exploring the aspiration space non-randomly. The trade-off between computation time and accuracy in approximation might also be studied in more detail; for example, it may be possible to be more selective in the number of iterations that are run for each particular case, by basing the decision on the number and diversity of the favourable aspiration vectors already obtained. Secondly one might attempt to employ the approach in a real-world experiment involving choices or orderings made by actual decision makers in order to investigate under which psychological and environmental conditions, if any, the proposed version of satisficing may be said to be an appropriate model of decision making.

References

[1] M. Allais. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Americaine. *Econometrica*, 21:503–546, 1953.

[2] V. Belton and T. Stewart. *Multiple criteria decision analysis: an integrated approach*. Kluwer Academic Publishers, Boston, 2002.

[3] W. Brock and S. Durlauf. A multinomial-choice model of neighborhood effects. *American Economic Review;*, 92(2):298–303, 2002.

[4] P. Chintagunta and J.-P. Dubé. Estimating a stockkeeping-unit-level brand choice model that combines household panel data and store data. *Journal of Marketing Research*, 42(August):368–379, 2005.

[5] I. Durbach. A simulation-based test of stochastic multicriteria acceptability analysis using achievement functions. *European Journal of Operational Research*, 170:923–934, 2006.

[6] I. Durbach and T. Stewart. Using expected values to simplify decision making under uncertainty. *Omega*, doi: 10.1016/j.omega.2007.02.001, 2007.

[7] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75:643–669, 1961.

[8] W.-T. Fu and W. Gray. Suboptimal tradeoffs in information seeking. *Cognitive Psychology*, 52:195–242, 2006.

[9] G. Gigerenzer and P. Todd. *Simple heuristics that make us smart*. Oxford University Press, New York, 1999.

[10] J. González-Pachón and C. Romero. Satisficing logic and goal programming: towards an axiomatic link. *INFOR*, 42(3):157–161, 2004.

[11] J. Hokkanen, R. Lahdelma, and P. Salminen. A multiple criteria decision model for analyzing and choosing among different development patterns for the helsinki cargo harbor. *Socio-Economic Planning Sciences*, 33(1):1–23, 1999.

[12] J. Ignizio. *Goal programming and extensions.* Lexington Books, Massachusetts, 1976.

[13] M. Isik and M. Khanna. Stochastic technology, risk preferences, and adoptions of site-specific technologies. *American Journal of Agricultural Economics*, 85(2):305–317, 2003.

[14] D. Kahneman and A. Tversky. Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

[15] W. Kakamura and R. Srivastava. An ideal-point probabilistic choice model for heterogenous preferences. *Marketing Science*, 5(3):199–218, 1986.

[16] R. Keeney and H. Raiffa. *Decisions with multiple objectives: preferences and value tradeoffs.* John Wiley & Sons, New York, 1976.

[17] R. Lahdelma, J. Hokkanen, and P. Salminen. SMAA – stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, 106:137–143, 1998.

[18] R. Lahdelma, S. Makkonen, and P. Salminen. Two ways to handle dependent uncertainties in multi-criteria decision problems. *Omega*, doi:10.1016/j.omega.2006.08.005, 2006.

[19] R. Lahdelma, K. Miettinen, and P. Salminen. Reference point approach for multiple decision makers. *European Journal of Operational Research*, 164(3):785–791, 2005.

[20] R. Lahdelma and P. Salminen. SMAA-2: stochastic multi-criteria acceptability analysis for group decision making. *Operations Research*, 49(3):444–454, 2001.

[21] J. Lee, K. Sudhir, and J. Steckel. A multiple ideal point model: capturing multiple preference effects from within an ideal point framework. *Journal of Marketing Research*, 39(Feb):73–86, 2002.

[22] G. Miller. The magical number seven, plus or minus two: some limits on our capacity of processing information. *Psychological Review*, 63:81–97, 1956.

[23] R. Pati, P. Vrat, and P. Kumar. A goal programming model for paper recycling system. *Omega*, 36(3):405–417, 2008.

[24] J. Payne, J. Bettman, and E. Johnson. *The adaptive decision maker*. Cambridge University Press, Cambridge, 1993.

[25] C. Romero. Extended lexicographic goal programming: a unifying approach. *Omega*, 29:63–71, 2001.

[26] R. Rust, K. Lemon, and V. Zeithaml. Return on marketing: using customer equity to focus marketing strategy. *Journal of Marketing*, 68(January):109–127, 2004.

[27] A. Saak and D. Henessy. Planting decisions and uncertain consumer acceptance of genetically modified crop varieties. *American Journal of Agricultural Economics*, 84(2):308–319, 2002.

[28] P. Seetharaman. Modeling multiple sources of state dependence in random utility models: a distributed lag approach. *Marketing Science*, 23(2):263–271, 2004.

[29] H. Simon. Rational choice and the structure of the environment. *Psychological Review*, 63:129–138, 1956.

[30] H. Simon. *The sciences of the artificial*. MIT Press, Cambridge, 3rd edition, 1996.

[31] V. Singh, K. Hansen, and S. Gupta. Modeling preferences for common attributes in multicategory brand choice. *Journal of Marketing Research*, 42(May):195–209, 2005.

[32] V. Srinivasan and A. Shocker. Linear programming techniques for the multidimensional analysis of preferences. *Psychometrika*, 38:337–369, 1973.

[33] T. Tervonen, H. Hakonen, and R. Lahdelma. Elevator planning stochastic multicriteria acceptability analysis. *Omega*, 36(3):352–362, 2008.

[34] T. Tervonen and R. Lahdelma. Implementing stochastic multicriteria acceptability analysis. *European Journal of Operational Research*, 178(3):500–513, 2007.

[35] J. Von Neumann and O. Morgernstern. *Theory of games and economic behavior*. Princeton University Press, 1944.