The use of the SMAA acceptability index in descriptive decision analysis

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

May 5, 2008

Abstract

This paper proposes and evaluates a descriptive multiattribute choice model based upon the notion of acceptability developed in the stochastic multicriteria acceptability analysis (SMAA) methods. The acceptability index is simply the relative proportion of all possible preference structures that support the selection of a particular alternative, and is related to a model of choice in which attribute evaluations are relatively stable but what is desired changes as if at random. The model is tested using two longitudinal surveys of FMCG markets in Europe. The acceptability index is found to be positively associated with relative purchase frequencies at the individual and aggregate level, inversely related to defection rates, and positively related to changes in relative purchase frequencies over a six-month period.

Keywords: choice theory; marketing; stochastic multicriteria acceptability analysis; consumer behaviour

1 Introduction

It sometimes seems that models of the way people should and do make choice decisions are surprisingly disparate. Multicriteria decision analysis (MCDA), the science of better decision-making, employs some methods that are also popular in the descriptive decision sciences – for example, multiattribute utility theory [e.g. 20] – but also uses other methods that have received little or no attention in the descriptive domain; for example, outranking methods [e.g. 40] and goal programming [e.g. 17]. Consumer behaviour and brand-choice models may make extensive use of utility-theoretic models [e.g. 31], but they also concern themselves with the construction of simple metrics that can be easily measured in survey questions and can shown to be correlated with actual purchase behaviour e.g. overall satisfaction in [29], intention to repurchase in [30], and various forms of 'commitment' in [39] and [13].

In this paper, output from a model developed and usually employed as a prescriptive decision aid is used as a simple measure of predicted relative purchase frequency; that is, as a model of descriptive choice behaviour. The measure is based upon the stochastic multicriteria acceptability analysis (SMAA) acceptability index and requires only an indication of which alternatives possess each of a set of attributes i.e. a traditional matrix of attribute evaluations that may be binary. This data is easily collected in survey form. Analytically, the model requires no parameter estimation and nothing more complex than Monte Carlo simulation is required. This makes it simple to implement, a distinct benefit because to a certain extent the lack of penetration of MCDA techniques into descriptive decision making can be explained by the difficulty of estimating the model parameters – for example, veto and other thresholds in the outranking methods, and aggregation metrics and aspiration levels in goal programming – without detailed and usually facilitated involvement on the part of the decision maker.

The measure is validated against various forms of purchase behaviour using datasets from two different fast-moving consumer goods (FMCG) markets in the United Kingdom and Spain. The paper makes two contributions to the decision analysis literature. From a prescriptive decision aiding perspective, the study results are the first empirical validation of stochastic multicriteria acceptability analysis in any real-world descriptive choice behaviour, and the paper thus extends the scope of these methods into that domain and in doing so attempts to strengthen the relationship between prescriptive and descriptive modelling. From a descriptive/brand-choice perspective, it proposes a measure of attitudinal equity that correlates well with purchase behaviour at the individual and aggregate level. It is also the only of the studied measures that is consistently associated with changes in relative purchase frequency over time once initial purchase frequencies are controlled for. That the "marketing" measure is constructed using techniques traditionally associated with operational research – multiattribute decision analysis and Monte Carlo simulation – further demonstrates the utility of integration between related fields of research.

The remainder of the paper is structured as follows. The following section provides a brief survey of the literature on relevant choice models in the descriptive and prescriptive decision sciences and introduces the SMAA family of methods. Section 3 introduces key notation, and section 4 provides the details of using SMAA as a model of buying behaviour. Section 5 outlines the details of the two longitudinal surveys used in the validation of the SMAA model. Sections 6 and 7 present the results obtained from the two studies – predominantly associations between the constructed measures and various forms of purchase behaviour – and discusses some of the implications of the results for the modelling of buying behaviour. Finally, section 8 provides some final conclusions and directions for future research.

2 Literature review

2.1 Review of relevant descriptive choice models

For the current paper it will be useful to distinguish between two broad kinds of descriptive brand-choice model. The first attempts to use a single measure of attitudinal equity as a proxy for the latent value of an alternative. Typical measures of aggregate attitudinal equity include overall satisfaction [29], affective commitment [39], repurchase intention [30], and overall 'efficacy' [30]. The popularity of these measures, particularly in commercial studies, derives from the parsimony of their measurement and the lack of a need for parameter estimation. Their ease of construction means that choice predictions can often be made directly from the single equity measure e.g. select the alternative with the highest overall satisfaction. Other examples of these models can be found in [5] and [21]

The second type of model tends to use multiple lower-level attributes such as marketing-mix variables or specific product or service characteristics such as price, quality, and so on. Such models generally provide a finer level of detail into the reasons for choice but also require considerably more effort in estimation. Furthermore, since different aspects may be important to different people, accommodating respondent heterogeneity using random-effects specifications becomes standard in these models [e.g. 34]. This means that in contrast to the models using a single aggregate measure of attitudinal equity, any choice prediction requires that substantial and often technically sophisticated model fitting is performed using some empirically-gathered dependent variable. Examples of these models can be found in [11] and [31].

The focus of this paper is on models that do not require any calibration against an empirically-gathered dependent variable. That is, I consider only models that construct a single measure of attitudinal equity from one or more survey questions without any parameter estimation. This excludes all so-called "proper" models of the second type. Proper linear models are those in which the weights of the input variables are estimated so as to fulfil some prespecified optimality condition, which requires a dependent variable. A model of the second type that can and will be considered however, is an "improper" linear model that assigns an equal weight to each attribute j. Dawes [3] has shown that predictions obtained from such an improper model can prove both accurate and robust to changes in external circumstance.

In considering the fundamental question of 'what causes a person to choose a particular alternative?', models that directly estimate a single quantity of aggregate attitudinal equity must be validated against observed choice behaviour. This behaviour has typically taken two forms: retention (the continuance of a relationship with a service or product) or share of wallet allocations (the percentage of money allocated by a customer to a particular service or product in a category). Many authors have shown that overall satisfaction has a positive effect on the intention to repurchase [e.g. 2, 22], retention [e.g. 12], and share of wallet [e.g. 21, 30]. The effect of satisfaction is generally of a moderate magnitude, with correlations between satisfaction and share of wallet and retention generally lying between 0.2 and 0.4 [e.g. 22, 21, 30]. However, other researchers have shown much smaller correlations between satisfaction and subsequent behaviour (e.g. correlations of 0.07 in [33] and 0.13 in [13]; also see [39]).

Other measures of attitudinal equity that have been proposed and tested against future buying behaviour include relationship quality [5], affective commitment [39, 14], calculative commitment [13], and attitudinal loyalty [41]. Although not strictly a measure of attitudinal equity, the effect of repurchase intentions on subsequent behaviour has also been studied repeatedly. Results have varied quite widely, with repurchase intentions by turn being shown to be substantially related ($r = 0.43$ in [18], $r = 0.44$ and 0.65 in [30]) and weakly related ($r = 0.24$ in [22] and $r = 0.11$ in [33]; also see [28]) to buying behaviour.

2.2 Review of relevant prescriptive choice models

Normative choice models are typically concerned with helping a decision maker or group of decision makers through a particular decision problem. Model parameters such as attribute weights and utility functions are not estimated in order to optimise fit to some external criteria but are decided on by the decision maker in consultation with a knowledgeable analyst so as to best represent the preferences of the decision maker(s). In some situations it may be difficult or impossible for a decision maker to explicitly give this preference information, for example in highly antagonistic political situations. The SMAA family of methods was developed in this context 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. That is, instead of asking 'which alternative is best given a particular set of preferences?', one asks 'what preferences might make this particular alternative the preferred one?'. 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 [e.g. 16] to complete assessment of preferences as for regular prescriptive decision aid [e.g. 19, 36].

The original SMAA method in [23] analyses the combinations of attribute weights that result in each of a set of prospective alternatives being selected when using an additive utility function. Subsequently, the SMAA methodology has been extended to several other prescriptive models of choice: to outranking in [35], to goal programming and reference point methods in [25] and [6], to data envelopment analysis in [27]; and to ordinal and mixed ordinal-cardinal evaluations in [24]. Practical examples of the application of SMAA in a prescriptive context are given in [19] and [36], while further technical details can be found in [37]. A recent study has proposed the use of SMAA in a descriptive setting to estimate the aspiration levels of a satisficing model [7]. Simulation studies conducted in that paper showed that the SMAA approach can perform well if strong performance is demanded on a small subset of the attributes. Several studies into the bounded rationality of decision making [e.g. 10, 15] suggest that this will often be the case for real-world decision problems. Nevertheless, the current paper is to my knowledge the first to apply SMAA to real-world descriptive decision making.

3 Notation

Reference is made to a set of I alternatives indexed by $i \in \{1, \ldots, I\}$, and a set of J attributes indexed by $j \in \{1, \ldots, J\}$ which are used as bases for evaluating the alternatives, so that x_{ij} denotes the evaluation of alternative i on attribute j. These evaluations can be gathered into an evaluation matrix $\mathbf{X}$. Attributes have been defined without loss of generality so that more is always preferred. In the validation studies presented later, all x_{ij} are binary i.e. $x_{ij} = 1$ if alternative i possesses attribute j and is zero otherwise. Alternatives are evaluated using an additive value function of the form

$$V_i = \sum_{j=1}^{J} w_j v_j(x_{ij}) \tag{1}$$

where $v_j(\cdot)$ is a function mapping attribute evaluations on attribute j to values and w_j is the weight attached to attribute j, to be interpreted in the usual swing weighting sense. A particular vector of attribute weights can be gathered into a weight vector $\mathbf{w}$, and the set of all possible weight vectors is denoted by $\mathbf{W}$. Where x_{ij} is binary, $v_j(\cdot)$ can be assumed to be linear so that $V_i = \sum_{j=1}^{J} w_j x_{ij}$. Alternatively, one may write $\mathbf{V} = \mathbf{wX}$.

4 Using SMAA to model buying behaviour

The use of SMAA as a means of modelling purchasing behaviour is motivated by a simple idea. Given a discrete set of non-dominated alternatives from which to choose and a matrix of evaluations $\mathbf{X}$, there must exist a combination of attribute weights $\mathbf{w}_i$ that makes each alternative i the preferred option i.e. $\mathbf{w}_i : V_i \geq V_k,\ \forall k = \{1, \ldots, I\}$. Such a weight vector is known in the SMAA literature as a *favourable weight vector*, with the set of all weight vectors that make alternative i the preferred one constituting the set of favourable weight vectors $\mathbf{W}_i$.

Since weights may be specified to an arbitrary precision, all non-dominated alternatives possess an infinite number of weights that would make them the preferred one. It therefore becomes necessary to use some summary measure of the set of favourable weights for each alternative. Several summary measures are possible, but the one that is most relevant to the current context as well as the most widely used generally [e.g. 19, 36] is the relative size of the hypervolume of favourable weight vectors

$$A_i = \frac{\mathrm{vol}(\mathbf{W}_i)}{\mathrm{vol}(\mathbf{W})} = \frac{\int_{\mathbf{W}_i} dw}{\int_{\mathbf{W}} dw} \tag{2}$$

The hypervolume A_i constitutes the so-called *acceptability index* because it gives an indication of the relative number of weight vectors that support the selection of each alternative. In the prescriptive decision aiding setting in which SMAA was developed and continues to be most widely used, the acceptability index can be used to identify and possibly exclude alternatives that are only supported by a very small volume of weight vectors – though of course even an alternative with a very small acceptability index may be preferred if one of those favourable weight vectors is the one actually held by the decision maker.

In a descriptive setting the acceptability index is used in a very similar way which leads directly to this paper's key hypotheses. This interpretation is based on the suggestion that an individual's preferences i.e. the relative weights allocated to each attribute, will show some variation over time, whether as a function of some internal change in what is desired, changes in purchase-to-purchase contextual factors or some other form of essentially random variation. If this variation was entirely random then the acceptability indices would represent the long-run purchase probability for each brand. Ehrenberg and his co-authors have shown in a series of books and papers [8, 11, 38] that purchase behaviour can be fruitfully modelled as occuring "in an *as if* random manner" at the individual level even if true randomness is only a feature of the model, which suggests that the acceptability index might be useful as a predictive measure of expected relative purchase frequencies, both at the individual and aggregate level. This is the first proposition that is tested in the current paper.

A related proposition is that the size of the hypervolume of favourable aspiration vectors might also be used to measure the extent to which the use of an alternative is robust to changes in the underlying preference structure. A particular shift in preferences is more likely to cause a change in the use of an alternative possessing a small acceptability index than one possessing a large index, simply because the region of favourable preferences is smaller. Put another way, the sensitivity of purchase behaviour to perturbations in the underlying weight vector possessed by an individual is likely to be greater for those alternatives with small acceptability indices relative to those with large acceptability indices, leading to greater defection rates and relatively greater decreases in spending. This constitutes the second and third of the research hypotheses below.

Hypothesis 1: *Relative purchase frequency is positively associated with the SMAA acceptability index, both at the individual and aggregate level.*

Hypothesis 2: *Probabilities of defection are higher from alternatives with relatively lower SMAA acceptability indices.*

Hypothesis 3: *Changes in relative purchase frequencies are positively associated with the*

SMAA acceptability index.

Two comments on the interpretation of SMAA described above seem necessary. Firstly, although it is founded on the sensitivity of behaviour to variation in preference structures, it is not necessary to describe the source of the variation and all perturbations of preferences are essentially treated as random. Of course any perturbation of preferences is unlikely to be entirely - or maybe even predominantly - random. Were this so, the acceptability index would represent the long-run selection probability of each alternative. The converse however - that a good fit between the acceptability index and observed selection probabilities implies randomly changing attribute weights - does not follow. One can say however, what effects might be observed if preference structures change as if at random. The same point has been made in [11] and [38].

A second point relates to practical implementation. In practice the set of favourable weight vectors $\mathbf{W}_i$ is approximated by a representative set of favourable weight vectors $\hat{\mathbf{W}}_i$ generated using Monte Carlo simulation. At each iteration, a random weight vector is generated and equation (1) is applied to arrive at a complete rank order. The simulated weight vector is simply added to the set of favourable weight vectors of the most-preferred alternative in the current rank order, and the process repeated. More generally, one can refer to a weight vector $\mathbf{w}_\Gamma$ that is consistent with some possibly incomplete rank order Γ. This allows one to extend the acceptability index to other positions in the rank order rather than just the first rank [26]. For example, it is possible to compute the set of weight vectors $\hat{\mathbf{W}}_i^k$ that make alternative i the k^{th}-ranked alternative and hence a "k^{th}-placed" acceptability index A_i^k. In the validation study to follow, a cumulative acceptability index $A_i^{(\kappa)} = \sum_{k=1}^{\kappa} A_i^k$ is used that can be interpreted as the probability that alternative i is contained in the first κ positions of the rank order implied by the underlying preferences.

The exact number of Monte Carlo iterations that are required to achieve a given level of estimation accuracy has been shown in [37] to depend on the quantity that is being estimated. To estimate the acceptability index within ξ of the true value with 95% confidence, one requires $1.96^2/4\xi^2$ iterations. For example, in order to achieve error limits of 0.01 on the estimated acceptability index, 9604 iterations are required.

5 Study design

In the following two sections the relationship between the SMAA acceptability index and buying behaviour is evaluated by testing the three research hypotheses described above using real-world longitudinal data obtained for two FMCG markets. The data has been kindly provided by Synovate and consists of two longitudinal studies conducted in different markets in different countries. The studies are of the toothpaste market in the United Kingdom (17 brands, $n_1 = 281$) and the laundry detergent market in Spain (11 brands, $n_2 = 361$). In both studies, respondents completed an initial survey at t_0 in which they identified

- $B_1(t_0)$: the set containing those brands the respondent was aware of
- $B_2(t_0)$: the set containing those brands the respondent had purchased in the last 3 months
- $B_3(t_0)$: the set containing those brands the respondent would consider buying, including those brands in $B_2(t_0)$

- Sat: overall satisfaction with each brand in B_1, evaluated on a 10-point scale from 1 ("terrible") to 10 ("perfect")
- PI: the intention to purchase each of the brands in B_3 within the next few weeks, evaluated on a 5-point scale from 0 ("definitely will not buy") to 1 ("definitely will buy")
- AdR: a binary indicator of whether or not the respondent could remember seeing any advertising for each of the brands in $B_3(t_0)$
- SoP(t_0): the relative proportion of the respondent's last 10 purchases allocated to each of the brands in $B_2(t_0)$

Respondents were also shown a list of relevant attributes and were asked to associate brands from $B_1(t_0)$ with each of the attributes. Respondents were explicitly told that they could select one, more than one, or no brand at all for each attribute statement. There were 12 and 13 attributes in the studies of the UK toothpaste and Spanish laundry detergent markets respectively. These responses were recoded to form the attribute evaluation matrix $\mathbf{X}$ with elements $x_{ij} = 1$ if brand i was selected as possessing attribute j and $x_{ij} = 0$ otherwise. In both datasets, one of the attributes was negatively worded ("quite expensive"). Responses on this attribute were simply coded to take on the value -1 if a brand possessed the negative attribute and zero otherwise i.e. preferences are increasing in all attributes. The x_{ij} form the input to the application of the SMAA procedure described in section 4, with the resulting output being an acceptability index $A_i^{(\kappa)}$ representing the relative likelihood in the absence of any further preference information that brand i appears in the first κ positions in a rank order of all brands. Different acceptability indices $A_i^{(\kappa)}$ for $\kappa = \{1, 2, 3, 4, 5\}$ were calculated on the basis of 20000 random weight vectors generated for each respondent, in line with results in [37]. Weights are generated so as to be uniformly distributed and sum to one. For the sake of brevity the section to follow only shows results for $\kappa = \{1, 3, 5\}$, with results for the excluded intermediate cases in almost all cases showing consistent pattern with those values that are shown.

In addition, a further measure of preference for each brand was constructed from a equally-weighted linear combination of the attribute values i.e. $\sum_j x_{ij}$. This modelled is labelled 'Dawes' after previous research by that author showing that this model can be an excellent and robust representation of preferences [3]. By setting all weights equal to one, the measure can be thought of as representing the net number of positive attributes associated with each brand. It is also consistent with the other measures in that it can be constructed from a simple set of survey questions without reference to actual or claimed purchase frequencies.

A follow-up survey of all respondents at t_1, approximately 6 months after the initial survey, was used to gather information on the set of brands the respondent had purchased in the 3 months prior to the t_1 survey i.e. $B_2(t_1)$, as well as the relative proportion of the respondent's previous 10 purchases attributed to each of the brands in $B_2(t_1)$ i.e. SoP(t_1). For ease of presentation, the analysis section to follow focuses on three target brands in each dataset – the brand with the highest penetration, a 'medium-sized' brand selected at random from those brands with penetrations from 20%–50%, and a 'small' brand selected at random from those brands with penetrations from 5%–10%. This approach avoids problems arising from the strong dependence conditions implied by the share of purchases by definition totalling 100% at t_0 and t_1 while allowing the consistency of any outcomes over brands to be considered. The main conclusions of the study are unchanged if all brands are considered. Zero and 11 dropouts were reported in the UK toothpaste and Spanish laundry detergent surveys respectively, leaving effective sample sizes of $n_1 = 281$ and $n_2 = 350$. In addition to the relative purchase frequencies described above, the initial and follow-up survey information was used to determine which respondents had defected from brands used at t_0 by the time of the follow-up survey at t_1.

6 Results

The results section is divided into three parts corresponding to each of the three research hypotheses identified above. First, the associations between the SMAA acceptability index and relative purchase frequencies at t_0 and t_1 are examined. Second, the acceptability indices of those who defect from their used brands are compared to those who remain with their brands. Finally, I consider whether changes in relative purchase frequencies are related to the acceptability index after controlling for initial purchase frequencies at t_0.

6.1 Associations between the acceptability index and relative purchase frequency

I begin by evaluating the performance of the acceptability index in the somewhat easier task of predicting aggregate market share before turning to associations at the individual level. Tables 1 and 2 show the relationship between predictions of aggregate purchase frequency i.e. a proxy for market share, obtained using five predictive measures and actual market share at t_0 and t_1, for all brands with market shares greater than 5%, in each market. Aggregated brand-level measures are obtained by simply averaging each measure over all respondents. For brevity only the best-performing acceptability index based only on the first position in the rank order is shown. All measures make use the relative rather than absolute performance levels i.e. relative satisfaction, advertising recall, net number of positive attributes, and purchase intentions have been used. This simply reflects a preference model in which purchase probabilities are proportional to absolute levels of satisfaction, advertising recall, net number of positive attributes, or purchase intentions respectively, and is necessary for market shares to sum to 100% at both an individual and aggregate level. The acceptability index is an inherently relative quantity and so does not require any standardisation. Finally, $\bar{\delta}(t_k) = \sum_{i=1}^{n} |\Psi_i - \mathrm{SoP}_i(t_k)|/n$ denotes the average absolute deviation between actual market share $\mathrm{SoP}_i(t_k)$ and predicted market share Ψ_i over all brands in each market. Significant differences between these values and the values of $\bar{\delta}(t_k)$ obtained using the acceptability index are flagged using a single, double, or triple asterisk superscript to denote significance at the 10%, 5% or 1% level respectively (the nonparametric Wilcoxin matched pairs test was used for this purpose since distributions of market shares are strongly skewed to the right).

Rank	SoP(t_0)	SoP(t_1)	Sat	AdR	Dawes	PI	$A_i^{(1)}$
1	44.8%	52.4%	28.7%	16.9%	30.3%	34.1%	46.2%
2	17.0%	13.7%	16.8%	14.1%	14.9%	16.1%	17.9%
3	9.9%	9.4%	9.9%	19.8%	9.3%	10.3%	7.7%
4	7.0%	8.6%	4.4%	2.6%	3.6%	5.4%	4.5%
5	5.9%	6.5%	11.5%	4.2%	9.0%	9.3%	5.8%
$\bar{\delta}(t_0)$			2.3%	4.4%***	2.4%***	1.6%	0.8%
$\bar{\delta}(t_1)$			3.3%	5.2%***	3.2%**	2.5%	1.5%

Table 1: Associations between survey measures and overall market share for all brands with market share greater than 5% in Dataset 1

The first-rank SMAA acceptability index offers very accurate predictions of aggregate relative purchase frequencies, particularly at t_0 but also at t_1. The average deviation between the market share predicted by the acceptability index and the actual market share amounts to 0.8% and 1.5% respectively. This is more accurate than any of the other measures, including purchase intentions themselves, although the difference is only significant at the 1% level for advertising recall and the relative net number of positive attributes. The latter result is particularly important because it shows that the acceptability index is demonstrably superior to the simpler-to-implement net count of the number of positive attributes. Another noteworthy result is that only the acceptability index is able to accurately predict the market share of the leading brand for this market, which is 44.8% at t_0 and 52.4% at t_1. The acceptability index for that brand is 46.2%, an absolute deviation of 1.4% and 6.2% respectively. The next most accurate measure is purchase intentions, which has absolute deviations of 10.7% and 18.3% respectively. The net number of positive attributes and overall satisfaction perform even worse, and advertising recall performs worst of all. The results in Table 1 thus provide a strong indication of the ability of the acceptability index to model buying behaviour at the aggregate level, particularly in relation to other existing measures.

Rank	SoP(t_0)	SoP(t_1)	Sat	AdR	Dawes	PI	$A_i^{(1)}$
1	30.1%	27.6%	25.1%	21.1%	26.7%	24.7%	30.9%
2	19.2%	21.5%	11.7%	1.8%	13.7%	16.1%	15.3%
3	11.8%	11.0%	14.2%	9.7%	12.8%	13.6%	11.1%
4	10.0%	9.2%	13.9%	11.7%	12.9%	11.6%	10.6%
5	6.3%	7.7%	7.9%	10.2%	7.7%	7.3%	7.1%
6	6.0%	6.7%	5.3%	4.8%	5.6%	5.5%	5.6%
$\bar{\delta}(t_0)$			2.4%**	5.4%***	1.7%	1.6%	1.1%
$\bar{\delta}(t_1)$			2.5%*	5.3%***	1.8%	1.8%	1.5%

Table 2: Associations between survey measures and overall market share for all brands with market share greater than 5% in Dataset 2

The conclusions that can be drawn from the second dataset do not differ materially from those of the first. That is, the acceptability index again gives the most accurate measure of market share at both t_0 and t_1, with average deviations of 1.1% and 1.5% respectively. This is better than any other measure although improvements only achieve statistical significance over overall satisfaction and advertising recall. The other measures all offer far better prediction of the share gained by the market leading brand than was the case for the first dataset. The estimate of the market leader's share provided by the relative net number of positive attributes and purchase intentions are in fact both slightly more accurate than the acceptability analysis at t_1, although they still tend to underestimate this share. Nevertheless, the results obtained from the second dataset support the claim that there is a strong association between aggregate market share and the acceptability index.

Having considered purchase behaviour in the aggregate, one can now turn to the considerably more difficult task of modelling individual buying behaviour. Tables 3 and 4 show correlations between each of the seven measures (three SMAA acceptability indices are now used, plus the four other measures) and the shares of purchases allocated to each brand at t_0 and t_1, for each market. Again, all measures use the relative rather than absolute performance levels. Significance at the 5% and 0.5% level is denoted using a single and double asterisk superscript respectively.

Outcome	Brand	Sat	AdR	Dawes	PI	$A_i^{(1)}$	$A_i^{(3)}$	$A_i^{(5)}$
SoP(t_0)	Large	0.65**	0.08	0.62**	0.85**	0.76**	0.61**	0.57**
	Medium	0.78**	0.48**	0.76**	0.88**	0.69**	0.58**	0.50**
	Small	0.61**	0.19**	0.67**	0.75**	0.68**	0.49**	0.45**
SoP(t_1)	Large	0.34**	0.03	0.42**	0.49**	0.52**	0.44**	0.39**
	Medium	0.63**	0.04	0.69**	0.72**	0.68**	0.47**	0.40**
	Small	0.15*	0.16*	0.16*	0.25**	0.34**	0.43**	0.43**

Table 3: Associations between survey measures and share of purchases in Dataset 1

All acceptability indices are significantly related to relative purchase frequency at the 0.5% level at both t_0 and t_1, providing strong support in favour of the first research hypothesis. The best performing form of the acceptability index $A_i^{(1)}$ achieves correlations of between 0.68 and 0.76 with share of purchases at t_0 and correlations of between 0.34 and 0.68 with share of purchases at t_1. With the exception of the small brand, the strength of association between both of the purchase frequencies and the acceptability indices decreases as more of the rank order is considered.

The first-rank acceptability index shows a higher correlation with relative purchase frequency at t_0 than either the net number of positive attributes or overall satisfaction for two of three brands, although all are markedly lower than the correlations obtained using purchase intentions. At t_1, correlations involving $A_i^{(1)}$ are again higher than the net number of positive attributes and overall satisfaction, this time for all three brands and to a greater extent than at t_0. In fact, the first-rank acceptability index exhibits higher correlations with relative purchase frequencies at t_1 than purchase intentions do for two of the three brands. Advertising recall is markedly more weakly associated with relative purchase frequency at both t_0 and t_1, and at t_1 is only significant for the small brand.

Outcome	Brand	Sat	AdR	Dawes	PI	$A_i^{(1)}$	$A_i^{(3)}$	$A_i^{(5)}$
SoP(t_0)	Large	0.62**	0.24**	0.69**	0.80**	0.67**	0.58**	0.55**
	Medium	0.65**	0.20**	0.76**	0.83**	0.73**	0.66**	0.65**
	Small	0.70**	0.18**	0.64**	0.75**	0.60**	0.58**	0.56**
SoP(t_1)	Large	0.34**	0.11*	0.43**	0.46**	0.43**	0.31**	0.30**
	Medium	0.33**	0.12*	0.50**	0.44**	0.49**	0.38**	0.37**
	Small	0.30**	0.09	0.34**	0.37**	0.41**	0.29**	0.28**

Table 4: Associations between survey measures and share of purchases in Dataset 2

The results obtained from the second dataset largely echo those obtained from the first. All acceptability indices are highly significantly related to relative purchase frequency at both t_0 and t_1, with the strength of associations declining as the number of ranks considered increases from one to three to five. Purchase intentions are most strongly correlated with relative purchase frequency at t_0 (with correlations between 0.75 and 0.83), but the first-rank acceptability index $A_i^{(1)}$ shows the highest correlations with purchase frequencies at t_1 for two of the three brands. Purchase intentions, the relative net number of positive attributes, and the first-rank acceptability index are all roughly equally strongly correlated with purchase frequency at t_1 (where correlations are between 0.34 and 0.50). Correlations involving overall satisfaction are comparable with the acceptability index and the net number of positive attributes at t_0 but are again lower at the later t_1.

6.2 Associations between the acceptability index and defection behaviour

Tables 5 and 6 show the relationship between each of the seven measures and the likelihood of full defection, for each market respectively. In these tables, the absolute rather than relative levels of satisfaction, advertisting recall, and net number of positive attributes have been used. Within each brand, the first two rows give the average value of each measure among non-defecting and defecting users respectively. The third row gives a Z-statistic obtained from a non-parametric Mann-Whitney test for equality of means. Results that are significant at the 5% level are superscripted with a single asterisk and results that significant at the 0.5% level are superscripted with a double asterisk. The fourth row shows the odds ratio (OR) associated with a logistic regression of full defection behaviour on each measure. The fifth and final row in each sub-table shows the Z-statistic associated with each odds ratio, and is superscripted to denote statistical significance in the same way as described above.

		n	Sat	AdR	Dawes	PI	$A_i^{(1)}$	$A_i^{(3)}$	$A_i^{(5)}$
Large	$\bar{X}_{\text{Stay}}$	181	8.28	0.33	7.83	0.82	0.65	0.81	0.84
brand	$\bar{X}_{\text{Defect}}$	16	7.81	0.13	3.63	0.56	0.18	0.36	0.43
	$Z_{\bar{X}}$		1.87	1.70	4.10**	3.80**	3.88**	3.95**	4.23**
	OR		0.73	0.29	0.71	0.03	0.05	0.08	0.10
	Z_{OR}		-1.65	-1.61	-4.13**	-3.59**	-3.59**	-3.99**	-3.85**
Medium	$\bar{X}_{\text{Stay}}$	36	8.58	0.50	6.50	0.83	0.42	0.67	0.71
brand	$\bar{X}_{\text{Defect}}$	23	8.26	0.39	4.48	0.65	0.11	0.39	0.47
	$Z_{\bar{X}}$		0.48	0.81	2.11*	2.74*	2.51*	2.28*	1.96*
	OR		0.61	0.64	0.84	0.08	0.04	0.24	0.32
	Z_{OR}		-1.42	-0.82	-2.08*	-2.36*	-2.72*	-2.30*	-1.92
Small	$\bar{X}_{\text{Stay}}$	14	7.79	0.14	5.36	0.73	0.18	0.39	0.47
brand	$\bar{X}_{\text{Defect}}$	15	7.00	0.07	3.87	0.47	0.08	0.24	0.30
	$Z_{\bar{X}}$		1.24	0.66	1.19	2.61*	1.22	0.92	0.95
	OR		0.69	0.43	0.89	0.02	0.30	0.40	0.41
	Z_{OR}		-1.31	-0.66	-1.08	-2.18*	-0.87	-1.00	-1.02

Table 5: Associations between survey measures and full defection behaviour in Dataset 1

The second research hypothesis, stating that probabilities of defection should be inversely related to the SMAA acceptability indices, is supported for two of the three brands considered in the first dataset but does not achieve statistical significance for the third and smallest considered brand. All relationships are in the hypothesised direction. That is, average acceptability indices for all brands are smaller among defecting users than among non-defecting ones, and odds ratios are all less than one, indicating a reduced risk of defection as the acceptability index increases. For the largest brand in the market, the acceptability index based only on the first rank $A_i^{(1)}$ indicates that on average 65% of all preferences supported the brand as 'best' among those who did not defect. In comparison just 18% of all preferences supported the brand as best among those who defected within the next six months. Similar results are obtained for the medium-sized brand and although results are not quite as significant as for the large brand they remain significant at the 5% level.

Significance levels are relatively constant regardless of the number of ranks considered by the acceptability index, and odds ratios indicate that those with acceptability indices around 0 (indicating a dominated alternative) are around 20 times more likely to defect than those with acceptability indices around 1 (indicating a dominating alternative). There is also evidence to suggest that the acceptability index $A_i^{(\kappa)}$ is sensitive to the size of a brand. For example, average values (irrespective of defection behaviour) of $A_i^{(1)}$ show a decline of 0.49 from largest to smallest brands (from 0.62 for the large brand to 0.30 and 0.13 for the medium and small brands respectively), while average values of $A_i^{(5)}$ decline by 0.43. The decline in the magnitude of the average acceptability indices as brand size gets smaller can be viewed as a manifestation of the 'double jeopardy' effect by which large brands not only have more users than small brands but also have the additional advantage that their users buy with greater frequency than users of small brands (see [9] for details).

The SMAA acceptability index compares favourably with the other attitudinal equity measures. Overall satisfaction is non-significant for all three brands. The Dawes measure is of roughly equal significance with the first-ranked acceptability index for all brands.

Purchase intentions are significant at the 0.5% level for the large brand and at the 5% level for both medium- and small-sized brands, and is thus the only measure that shows a consistently significant relationship with full defection behaviour. Whether or not any advertising has been seen for a brand has no significant association with any defection behaviour, although the results are in the hypothesised direction i.e. fewer defecting users claim to have seen advertising than those not defecting.

		n	Sat	AdR	Dawes	PI	$A_i^{(1)}$	$A_i^{(3)}$	$A_i^{(5)}$
Large	$\bar{X}_{\text{Stay}}$	116	7.89	0.63	5.40	0.79	0.54	0.82	0.88
brand	$\bar{X}_{\text{Defect}}$	39	4.79	0.46	2.95	0.30	0.17	0.52	0.55
	$Z_{\bar{X}}$		4.70**	1.83	4.31**	6.60**	5.40**	2.87**	4.30**
	OR		0.63	0.5	0.74	0.01	0.04	0.17	0.14
	Z_{OR}		-4.94**	-1.83	-4.09**	-6.17**	-4.59**	-3.93**	-4.34**
Medium	$\bar{X}_{\text{Stay}}$	44	7.57	0.45	4.61	0.73	0.44	0.72	0.79
brand	$\bar{X}_{\text{Defect}}$	34	3.26	0.50	1.56	0.20	0.07	0.23	0.29
	$Z_{\bar{X}}$		5.15**	-0.40	4.69**	5.92**	4.94**	4.64**	4.18**
	OR		0.56	1.2	0.63	0.01	0.02	0.07	0.08
	Z_{OR}		-4.05**	0.40	-3.81**	-4.85**	-3.43**	-4.36**	-4.32**
Small	$\bar{X}_{\text{Stay}}$	13	7.23	0.69	3.85	0.79	0.28	0.63	0.70
brand	$\bar{X}_{\text{Defect}}$	14	1.71	0.57	1.29	0.13	0.12	0.19	0.25
	$Z_{\bar{X}}$		3.26**	0.64	3.04**	4.23**	2.29*	2.47*	2.14*
	OR		0.56	0.59	0.70	0.001	0.19	0.07	0.10
	Z_{OR}		-2.96**	-0.65	-2.08*	-2.92**	-1.23	-2.44*	-2.38*

Table 6: Associations between survey measures and full defection behaviour in Dataset 2

The results for the second dataset reinforce the main conclusions drawn from the first. That is, there is again strong support for the second research hypothesis that the SMAA acceptability index is inversely related to defection likelihood. All three SMAA acceptability indices show some statistical significance for all three brands considered, although there are some discrepancies in which index is most significant in each case. The first-rank index $A_i^{(1)}$ is highly significant for the large- and medium-sized brands but only significant at the 5% for the small brand (with the odds ratio not significant at all in this case, although the relationship is in the hypothesised direction). The other two acceptability indices are both significant under all conditions. The results again show that the average SMAA acceptability index among users of large brands tends to be greater than among users of small brands.

The main difference in results is the better performance of overall satisfaction, which is consistently highly significant for all three considered brands. The Dawes measure is again of roughly equal significance with the first-ranked acceptability index for all brands, and purchase intention is again highly significantly associated with future defection behaviour. The average stated likelihood of repurchase is also highly consistent across brand sizes among those not defecting. Although this likelihood appears to decline among defecting users as brand size decreases, the change is not significant (Kruskal-Wallis $H = 1.54$, $p = 0.45$). This is consistent with the results obtained using the other dataset, as is the finding that advertising recall has no significant association with any defection behaviour.

6.3 Associations between the acceptability index and changes in relative purchase frequency

Tables 7 and 8 show the partial correlations between the seven measures with relative purchase frequencies at t_1, after controlling for the effect of purchase share at t_0. It is also possible to test hypothesis 3 by regressing changes in purchase frequency i.e. $\text{SOP}(t_1) - \text{SOP}(t_0)$, on both relative purchase frequency at t_0 and each of the predictive measures. The conclusions to be drawn are the same using either analysis but the partial correlations more concisely address the question of whether an association exists between a measure and changes in relative purchase frequencies over time. As before, significance at the 5% and 0.5% level is denoted using a single and double asterisk superscript respectively.

Outcome	Brand	Sat	AdR	Dawes	PI	$A_i^{(1)}$	$A_i^{(3)}$	$A_i^{(5)}$
$\text{SoP}(t_1\|t_0)$	Large	-0.06	-0.02	0.09	0.01	0.17**	0.15*	0.15*
	Medium	-0.02	-0.21**	0.18**	-0.01	0.29**	0.00	0.00
	Small	-0.16*	0.01	-0.19**	-0.13*	0.06	0.27**	0.29**

Table 7: Associations between survey measures and share of purchases in Dataset 1

The acceptability index is the only measure that exhibits reasonably consistent significant correlations with changes in relative purchase frequencies over time. Although correlations are not very high, the correlations obtained from the linear form of $A_i^{(1)}$ is positive and significantly different from zero at the 0.5% level for two of the three brands. Although the correlation for the small brand is not significant, it is in the intended direction. Correlations for the two other acceptability indices are also significant for two of the three brands. These correlations provide some support in favour of the third research hypothesis, to the effect that changes in the relative purchase frequency are positively related with the SMAA acceptability index. No other measure approaches a similar level of consistent statistical significance. The net number of positive attributes is highly significant for the medium-size brand and is significant but in the wrong direction for the small brand. The correlations obtained using purchase intentions are either not significant or in the wrong direction.

Outcome	Brand	Sat	AdR	Dawes	PI	$A_i^{(1)}$	$A_i^{(3)}$	$A_i^{(5)}$
$\text{SoP}(t_1\|t_0)$	Large	-0.04	-0.03	0.05	-0.02	0.06	-0.04	-0.04
	Medium	-0.04	0.01	0.15**	-0.04	0.15**	0.02	0.02
	Small	-0.06	0.01	0.04	0.02	0.17**	0.01	0.00

Table 8: Associations between survey measures and share of purchases in Dataset 2

The results obtained from the second dataset are again consistent with those obtained from the first. The acceptability index $A_i^{(1)}$ shows a significant partial correlation with relative purchase frequencies at t_1 for two of the three brands and is notably more significant than any of the other measures, achieving correlations of 0.15 and 0.17 for the medium and small brand respectively. Again, the relative net number of positive attributes exhibits significant correlations for the medium-sized brand only. There is thus again evidence supporting the hypothesised positive relationship between changes in the relative purchase frequency and the first-ranked SMAA acceptability index for two of the three brands and for the relative advantage of this measure over the others in this regard.

7 Discussion

The results presented in the previous section provide support for the three research hypotheses in section 4. The aim of this section is to examine some of the implications of these results for research into consumer decision making at a theoretical and practical level. Section 4 outlined how the SMAA acceptability index is constructed by randomly generating vectors in the weight space and computing the relative volume that is compatible with a particular partial rank ordering of alternatives. The first-ranked acceptability index $A_i^{(1)}$ for example, is simply the proportion of all possible weight vectors that would lead to the selection of alternative i. While a 'randomly fluctuating preferences' model is by no means the only model capable of generating significant correlations between the acceptability index and purchase frequencies, it is of interest that the acceptability index results suggest a model of consumer decision making that gives the *appearance* of random choice given the set of attribute evaluations.

Previous studies have found a similarly strong empirical relationship between actual purchase frequencies and ones predicted by a model that assumes that a consumer's choices among the available brands follows a multinomial distribution [11]. In that model, consumer choices also occur in an "*as if* random manner" given a set of initial purchase probabilities distributed over the set of available brands. This bears a clear resemblance to the SMAA model, which posits that choices occur in a similar "as if random" manner given the initial attribute evaluations and an additive value function. The SMAA model therefore provides a plausible explanation for how the initial purchase probabilities in [11] might be generated i.e. from a consideration of the attribute evaluations. Furthermore, it provides an empirical basis for estimating these probabilities from attitudinal survey data – the model in [11] requires that the purchase probabilities be estimated from behavioural data, either in the form of detailed longitudinal purchase data or average purchase rates.

It is also instructive to compare the first-rank SMAA acceptability index with three other models based on attitudinal survey data – overall satisfaction, the relative net number of positive attributes, and acceptability indices including more than the first rank. A model of choice based on satisfaction suggests that a consumer's propensity to purchase a brand will be some function of an overall evaluation of his or her satisfaction with the brand. The generally poorer associations of both absolute and relative satisfaction with actual purchase frequencies in comparison to both the acceptability index and the Dawes measure suggests that overall evaluations of satisfaction do not perform as well as a set of attribute evaluations in explaining purchasing behaviour[1]. This point has been made previously [e.g. 4, p. 393].

A model of choice based on the net number of positive attributes a brand possesses suggests a model in which consumers form opinions about brands by tallying up pros and cons, regardless of whether other brands also possess those same features. The main difference between this model and the one implied by the use of an acceptability index is

[1] One might make the point that overall satisfaction mediates the relationship between the attributes and final purchase behaviour as done, for example, in the return on quality (ROQ) model proposed by [32]. According to [1], mediation requires that (a) satisfaction influence purchase behaviour, (b) the acceptability index influences satisfaction, and (c) the effect of the acceptability index on purchase behaviour is weakened or eliminated when satisfaction is used as a covariate. All three of these conditions hold in all regressions involving relative purchase frequency at t_0 and t_1 but reductions in the significance of the acceptability index when including satisfaction as a covariate are in general limited, and satisfaction can at most be said to partially mediate the relationship between the acceptability index and purchase behaviour.

that the former allows for the selection of dominated alternatives while the latter does not. Although the results are often fairly close, the first-rank acceptability index notably outperforms the relative net number of positive attributes in two areas: in the prediction of aggregate market share, and in the association with changes in relative purchase frequencies after taking into account purchase frequencies at t_0. This suggests that precluding the selection of dominated alternatives is a useful feature of a model of buying behaviour and that the acceptability index is a worthwhile extension to the simpler equally-weighted linear model.

A more speculative comparison can also be made between the first-rank acceptability index and those based on more than the first position in the rank order. Since the acceptability index $A_i^{(\kappa)}$ gives the probability that alternative i appears in the first κ positions in the rank order, a choice strategy based on $A_i^{(\kappa)}$ would consist of a random choice between the best κ alternatives. For $\kappa = 1$ this means that the best alternative is always chosen, and for $\kappa = n$ choice is completely at random. For intermediate values of κ, a two-stage process is implied in which alternatives are first rank ordered from best to worst according to (1) and the worst $n-\kappa$ alternatives eliminated; and then a random choice is made from among the remaining alternatives. The fact that correlations with relative purchase frequencies decline with increases in κ suggests that this additional complexity is unwarranted.

8 Conclusions and future research

In this paper I have proposed and evaluated the use of stochastic multicriteria acceptability analysis (SMAA) as a model of purchase behaviour and as a means of identifying those customers at risk of either full defection or reducing their relative purchase frequency. The use of SMAA is based on an operationally simple Monte Carlo simulation in which sets of attribute weights are randomly generated and the resultant preference orderings of alternatives observed. The acceptability index is the relative proportion of all generated weight vectors that lead to a particular partial rank order. The most widely used version of the acceptability index is simply the relative proportion of all generated weight vectors that led to the selection of each alternative as 'best', but two further indices were also tested: those relating to weight vectors in which an alternative appeared anywhere in the top three ranks or top five ranks respectively. Use of the acceptability index is consistent with the notion that attribute evaluations i.e. what people think of an alternative, are relatively stable but that attribute weights i.e. what people want, are relatively unstable and may change from purchase to purchase. This change is likely to be dependent on many subtle situational factors including environmental conditions and decision maker psychology, but within SMAA it is modelled as if it were random.

The ability of the SMAA acceptability index to predict buying behaviour was tested using two longitudinal FMCG datasets. The hypotheses tested were that relative purchase frequency would be positively associated with the acceptability index, that probabilities of defection would be higher from those alternatives with relatively lower acceptability indices, and that changes in the relative amount of spending allocated to an alternative would be positively related to that alternative's acceptability index. All hypotheses were supported for the majority brands of considered, and results that did not achieve statistical significance were nevertheless in the hypothesised direction. The following key results emerged from the validation study:

1. The SMAA acceptability index is fairly strongly positively correlated with relative

purchase frequencies a short period into the future. Results were significant at the 0.5% level for all six brands with correlations generally between 0.4 and 0.6.

2. Defecting users have on average substantially lower acceptability indices than non-defecting users, indicating that the alternative is supported by fewer weight vectors i.e. in fewer situational contexts, in the defecting group. Results were significant at the 0.5% level for five brands but non-significant for the sixth.

3. Changes in relative purchase frequencies are positively associated with the first-rank acceptability index, after controlling for relative purchase frequencies at the initial time point. Although correlations are not large (between 0.06 and 0.29), they are significant at the 0.5% level for four of six brands.

4. Results obtained using the first-ranked acceptability index are generally superior to those obtained using an acceptability index taking into account a greater proportion of the rank order.

5. Results obtained using the first-ranked acceptability index are demonstrably better than those obtained using either overall satisfaction or the equally-weighted linear model. The former suggests that lower-level attributes should be taken into account (rather than a single overall measure of satisfaction) while the latter suggests that there are benefits in using the more comprehensive model implied by SMAA.

The strong results obtained for the first-rank acceptability analysis make two main contributions. Firstly, when taken as a whole they support and extend the results in [7] suggesting that SMAA may be useful in descriptive decision-making contexts as well as in the prescriptive ones that they were developed for and are more traditionally applied to. The current paper applies the SMAA method to the specific context of predicting customer defections and/or relative purchase frequencies. Results indicate that SMAA can provide predictive measures that are at least as powerful as some of the survey measures currently used, including direct statements about intended future behaviour. Whereas the results in [7] were based upon a simulation study of hypothetical decision making, the current paper analyses real-world choice behaviour by making use of two real-world longitudinal studies and is thus in a sense a stronger validation of the descriptive usefulness of SMAA. Secondly, the results join previous research on the appearance of randomness in brand choice reported in [11] and [38] and provide a possible mechanism for how the purchase probabilities that are the inputs to multinomial brand choice in that model are generated. They also have the practical advantage of being computable from easily-administered attitudinal survey questions.

The main aim of the current paper has been to provide an empirical validation of SMAA as a descriptive decision analysis tool; that is, using a model of choice in which preferences change as if at random. The current model makes no attempt to locate the current preference structure i.e. the currently held underlying weight vector, but in considering directions for future research one possibility is to attempt to identify these preferences (using the estimation approach in [7], for example) and to limit exploration of the weight space to some vicinity of the currently-held weights. Such a process would offer scope to limit the extent that perturbations are 'random' but is limited in its ability to model situations in which strongly different preference structures are possible. Further applications of the acceptability index, particularly to environments where interpurchase times are longer than the FMCG markets considered in this paper, would also be worthwhile.

Acknowledgements

I would like to gratefully acknowledge Synovate and in particular Dr. Jan Hofmeyr for providing me with the two datasets.

References

[1] R. Baron and D. Kenny. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182, 1986.

[2] R. Bolton and J. Drew. A longitudinal analysis of the impact of service changes on customer attitudes. *Journal of Marketing*, 55(1):1–9, 1991.

[3] R. Dawes. The robust beauty of improper linear models in decision making. *American Psychologist*, 34(7):571–582, 1979.

[4] R. Dawes. Proper and improper linear models. In T. Connolly, K. Hammond, and H. Arkes, editors, *Judgment and decision making: an interdisciplinary reader*, chapter 23. Cambridge University Press, 2000.

[5] K. De Wulf, G. Odekerken-Schröder, and D. Iacobucci. Investments in consumer relationships: a cross-country and cross-industry exploration. *Journal of Marketing*, 65(4):33–50, 2001.

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

[7] I. Durbach. On the estimation of a satisficing model of choice using stochastic multicriteria acceptability analysis. *Omega*, doi: 10.1016/j.omega.2007.09.001, 2007.

[8] A. Ehrenberg. *Repeat-buying: facts, theory and applications*. Charles Griffin & Company Limited, London, 1988.

[9] A. Ehrenberg, G. Goodhardt, and T. Barwise. Double jeopardy revisited. *Journal of Marketing*, 54(3):82–91, 1990.

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

[11] G. Goodhardt, A. Ehrenberg, and C. Chatfield. The Dirichlet: a comprehensive model of buying behaviour. *Journal of the Royal Statistical Society. Series A (General)*, 147(5):621–655, 1984.

[12] J. Goodman and D. Ward. The importance of customer satisfaction. *Direct Marketing*, 56(8):23–6, 1993.

[13] A. Gustafsson, M. Johnson, and I. Roos. The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention. *Journal of Marketing*, 69(4):210–218, 2005.

[14] J. Hofmeyr and B. Rice. *Commitment-led marketing: the key to brand profits is in the customer's mind*. Wiley, 2000.

[15] R. Hogarth and N. Karelaia. Simple models for multiattribute choice with many alternatives: when it does and does not pay to face trade-offs with binary attributes. *Management Science*, 51(12):1860–1872, 2005.

[16] 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.

[17] J. Ignizio and C. Romero. Goal programming. In H. Bidgoli, editor, *Encyclopedia of information systems*, volume 2, pages 489–499. Elsevier, USA, 2003.

[18] F. Juster. Consumer buying intentions and purchase probability: an experiment in survey design. *Journal of the American Statistical Association*, 61(315):658–696, 1966.

[19] J. Kangas, J. Hokkanen, A. Kangas, R. Lahdelma, and P. Salminen. Applying stochastic multicriteria acceptability analysis to forest ecosystem management with both cardinal and ordinal criteria. *Forest Science*, 49(6):928–937, 2003.

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

[21] T. Keiningham, T. Perkins-Munn, and H. Evans. The impact of customer satisfaction on share-of-wallet in a business-to-business environment. *Journal of Service Research*, 6(1):37, 2003.

[22] P. LaBarbera and D. Mazursky. A longitudinal assessment of consumer satisfaction/dissatisfaction: the dynamic aspect of the cognitive process. *Journal of Marketing Research*, 20(4):393–404, 1983.

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

[24] R. Lahdelma, K. Miettinen, and P. Salminen. Ordinal criteria in stochastic multicriteria acceptability analysis (SMAA). *European Journal of Operational Research*, 147(1):117–127, 2003.

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

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

[27] R. Lahdelma and P. Salminen. Stochastic multicriteria acceptability analysis using the data envelopment model. *European Journal of Operational Research*, 170(1):241–252, 2006.

[28] V. Morwitz, E. Johnson, and D. Schmittlein. Does measuring intent change behavior? *The Journal of Consumer Research*, 20(1):46–61, 1993.

[29] R. Oliver. A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, 17(4):460–469, 1980.

[30] T. Perkins-Munn, L. Aksoy, T. Keiningham, and D. Estrin. Actual purchase as a proxy for share of wallet. *Journal of Service Research*, 7(3):245, 2005.

[31] R. Rust, K. Lemon, and V. Zeithaml. Return on marketing: using customer equity to focus marketing strategy. *Journal of Marketing*, 68(1):109–27, 2004.

[32] R. Rust, A. Zahorik, and T. Keiningham. Return on quality (ROQ): making service quality financially accountable. *Journal of Marketing*, 59(2):58–70, 1995.

[33] K. Seiders, G. Voss, D. Grewal, and A. Godfrey. Do satisfied customers buy more? Examining moderating influences in a retailing context. *Journal of Marketing*, 69(4):26–43, 2005.

[34] V. Singh, K. Hansen, and S. Gupta. Modeling preferences for common attributes in multi-category brand choice. *Journal of Marketing Research*, 42(2):195–209, 2005.

[35] T. Tervonen, J. Figueira, R. Lahdelma, J. Dias, and P. Salminen. A stochastic method for robustness analysis in sorting problems. *European Journal of Operational Research*, doi:10.1016/j.ejor.2007.09.008, 2007.

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

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

[38] M. Uncles, A. Ehrenberg, and K. Hammond. Patterns of buyer behavior: regularities, models, and extensions. *Marketing Science*, 14(3):71–78, 1995.

[39] P. Verhoef. Understanding the effect of customer relationship management efforts on customer retention and customer share development. *Journal of Marketing*, 67(4):30–45, 2003.

[40] P. Vincke. Outranking approach. In T. Gal, T. Stewart, and T. Hanne, editors, *Multicriteria decision making: advances in MCDM models, algorithms, theory, and applications*, chapter 11. Kluwer Academic Publishers, Dordrecht, 1999.

[41] J. Wirtz, A. Mattila, and M. Oo Lwin. How effective are loyalty reward programs in driving share of wallet? *Journal of Service Research*, 9(4):327–334, 2007.