Boundary linear utility and sensitivity of decisions with imprecise utility trade-off parameters

Malcolm Farrow
Newcastle University, UK
Malcolm.Farrow@newcastle.ac.uk

Michael Goldstein
Durham University, U.K.
Michael.Goldstein@durham.ac.uk

Abstract

In earlier work we have developed methods for analysing decision problems based on multi-attribute utility hierarchies, structured by mutual utility independence, which are not precisely specified due to unwillingness or inability of an individual or group to agree on precise values for the trade-offs between the various attributes. Our analysis is based on whatever limited collection of preferences we may assert between attribute collections. In this paper we show how to assess the robustness of our selected decision using the properties of boundary linear utility.

Keywords. Robust decisions, imprecise utilities, utility hierarchies, mutual utility independence, boundary linear utility, sensitivity analysis.

1 Introduction

In two earlier papers we have developed a methodology for decision analysis with multi-attribute utilities which does not require the specification of precise trade-offs between different risks. Multi-attribute utilities may be imprecisely specified, due to an unwillingness or inability on the part of a client to specify fixed risk trade-offs or because of disagreement within a group with responsibility for the decision.

In [3] we introduced our approach to constructing imprecise multi-attribute utility hierarchies and finding the Pareto optimal rules. We described the structure which we use, which is based on a utility hierarchy with utility independence at each node, explained the notion of imprecise utility trade-offs for such a hierarchy, based on limited collections of stated preferences between outcomes, and used Pareto optimality, over the set of possible trade-off specifications, to reduce the set of alternatives. These methods and some associated theory are summarised in Section 2 of this paper.

We are particularly concerned with problems where the number of alternatives among which we must choose is large. Many real decision problems, for example in experimental design, have very large spaces of possible choices. Relaxing the requirement for precise trade-off specification reduces our ability to eliminate rules, i.e. choices, by dominance and can leave us with a large class of rules, none of which is dominated by any other over the whole range of possible trade-offs allowed by the imprecise specification. We are therefore faced with the need for practical ways to reduce the decision space which are tractable even when the decision space is very large and there is a complicated multi-attribute utility structure to consider. In [4] we described ways to reduce further the class of alternatives that we must consider, by eliminating rules which are "ε-dominated" and combining rules which are "ε-equivalent." We explored the effects of different values of ε and of different parts of the hierarchy to see when and why rules are eliminated.

To choose a single rule d^* from our reduced list, we can use the boundary linear utility approach described in [3], or choose the rule which is the last to be eliminated as we increase the value of our ε criterion as described in [4]. We can then find the set D^* of rules which are "almost equivalent" to d^* and perhaps use secondary considerations to choose among them. We review boundary linear utility in Section 3 of this paper.

In Section 4 we describe methods, based on the boundary linear utility, for exploring the sensitivity of possible choices to variation in the utility trade-offs. This helps us to find a decision which, as far as possible, is a good choice over the whole range of possible trade-offs.

The practical implementation of our approach is illustrated throughout by an example concerning the introduction of a new course module at a university, which we first described in [4].

2 Mutually utility independent hierarchies and imprecise utility tradeoffs

2.1 Mutually utility independent hierarchies

In [3] we proposed a general class of multi-attribute utility functions. This uses the concept of mutual utility independence among sets of attributes in order to impose a structure on the utility function. Attributes $\underline{Y} = (Y_1, ..., Y_k)$ are *utility independent* of the attributes $\underline{Z} = (Z_1, ..., Z_r)$ if conditional preferences over lotteries with differing values of $\underline{Y}$ but fixed values, $\underline{z}$, of $\underline{Z}$, do not depend on the particular choice of $\underline{z}$. Attributes $\underline{X} = (X_1, ..., X_s)$ are *mutually utility independent* if every subset of $\underline{X}$ is utility independent of its complement. If attributes $\underline{X}$ are mutually utility independent, then the utility function for $\underline{X}$ must be given by the *multiplicative form*

$$U(\underline{X}) = B^{-1}\left\{\prod_{i=1}^{s}[1 + ka_iU_i(X_i)] - 1\right\}, \qquad (1)$$

where B does not depend on $U_1(X_1), \ldots, U_s(X_s)$, or the *additive form*

$$U(\underline{X}) = \sum_{i=1}^{s} a_iU_i(X_i), \qquad (2)$$

(see [6]) where $\mathrm{U}_i(\mathrm{X}_i)$ is a conditional utility function for attribute X_i, namely an evaluation of the utility of X_i for fixed values of the other attributes. The coefficients in (1) and (2) are the *trade-off parameters*; the a_i reflect the relative importance of the attributes and k reflects the degree to which rewards may be regarded as complementary, if $k > 0$, or as substitutes, if $k < 0$.

The assumption of mutual utility independence, which many people would often be prepared to make, is enough in itself to reduce the problem to one of considering a finite number of parameters.

Keeney and Raiffa [6] also describe the idea of a hierarchy of utilities, as follows. We form an overall multi-attribute utility from marginal utilities for the various attributes by a hierarchical structure in which, at each node, several utilities are merged into a combined utility. This combined utility is merged with others at a node in the next level until, finally, one overall utility function is formed. If, at each node, we have mutual utility independence for the utilities combined at that node, then we term such a utility function a *Mutually Utility Independent Hierarchic (MUIH)* utility. Thus, in a MUIH utility, at each node we combine utilities using either (1) or (2).

Our hierarchical structure allows us to relax the requirement for overall mutual utility independence by allowing the user to specify utility independence just at the nodes of the hierarchy and, of course, the user can choose this structure.

In our utility hierarchy we consider the overall utility node to be at the "top" level and the predecessors of a node to be at "lower" levels. We refer to the nodes corresponding to the individual attributes, that is nodes which have no predecessors, as *marginal nodes*. We refer to a direct predecessor of a node as a *parent* and a direct successor as a *child*. For each node n, we denote by $H(n)$, the *sub-hierarchy* under n, where $H(n)$ is the set of nodes containing n and all of its predecessors. We divide the child nodes in the hierarchy into the following three types:

1. an *additive node*, where utilities are combined as in (2) with $\sum_{i=1}^{s} a_i \equiv 1$ and $a_i > 0$ for $i = 1, \ldots, s$;

2. a *binary node*, where precisely two utilities are combined, where we rescale the combined utility as
$$U = a_1U_1 + a_2U_2 + hU_1U_2 \qquad (3)$$
where $0 < a_i < 1$ and $-a_i \leq h \leq 1 - a_i$, for $i = 1, 2$, and $a_1 + a_2 + h \equiv 1$. Note that (3) is derived by setting $s = 2$ and $h = ka_1a_2$ in (1).

3. a *multiplicative node*, where more than two utilities are combined and the parameter k in (1) may be nonzero. We scale the utility using
$$B = \prod_{i=1}^{s}(1 + ka_i) - 1 \qquad (4)$$
with $a_1 \equiv 1, k > -1$ and, for $i = 1, \ldots, s$, we have $a_i > 0$ and $ka_i > -1$.

For each child node n, we denote by $\underline{\phi}_n = (\phi_{n,1}, \ldots, \phi_{n,m(n)})$ the collection of trade-off parameters which determine how the parent utilities at node n are combined to give the value at the child node. Thus, each $\phi_{n,j}$ corresponds to an a_i in (2) an a_i or h term in (3), or an a_i or k in (1). If there are N child nodes, then we denote by $\underline{\theta} = (\underline{\phi}_1, \ldots, \underline{\phi}_N)$ the collection of all the trade-off parameters in the hierarchy. If we allow imprecision in some of the elements of $\underline{\theta}$, then we refer to the resulting utility specification as an *imprecise independence hierarchy (IIH)*. If the hierarchy contains only additive and binary nodes, then we refer to the specification as a *simple imprecise independence hierarchy (SIIH)*

The utility at each child node is determined both by the values of the utilities at the marginal nodes and

also by the choice of trade-off parameters. As we shall vary the trade-off parameters, and thus the utilities at the child nodes, we require a standard scale for all utilities in the IIH, whose interpretation does not depend on the choice of trade-off parameters. This is constructed as follows.

As the marginal utility at each marginal node is expressed in a utility scale, we norm all the marginal utilities to lie between 0, the worst outcome that we shall consider for the problem, and 1, the best outcome. The effect of the scalings that we have chosen for additive, binary and multiplicative nodes is that, at each node n in the hierarchy, the utilities of C_n and c_n are 1 and 0 respectively, where C_n is an outcome such that all marginal predecessor nodes have utility 1, and c_n is an outcome such that all marginal predecessor nodes have utility 0. Therefore, a utility value of u at node n may always be interpreted as the utility of a gamble giving C_n with probability u and c_n with probability $1-u$, irrespective of the chain of trade-off parameters in the hierarchy. This utility scale is termed the *standard scale* for the hierarchy. Throughout this paper, all utilities are assumed to be on the standard scale.

2.2 Example: Designing a new course module at a university

In [4] we introduced an example concerning the design of a new course module at a university. We use the same example here to illustrate our approach. The module is to contain six units, or topics, each of which may, for the purpose of this example, be considered to be of the same size in the sense that, given the same teaching method, they would require the same length of time. Each topic could be taught by any one of three teaching methods, denoted as follows:

Lect : a traditional course of lectures and tutorials.

Lab : a laboratory-based course using a computer algebra package.

OL : an "open learning" course without lectures or formal laboratory sessions.

Thus we have $3^6 = 729$ possible choices of combinations of teaching methods. We can denote a choice $(\mu_1, \ldots, \mu_6)$ where $\mu_i = 1$,2 or 3 according to which method is used for unit i. (In practice there are additional choices to be made, but we do not wish to introduce unnecessary complexity into this example). The attributes which we consider in our analysis are as follows. Further details are given in [4].

- For students:

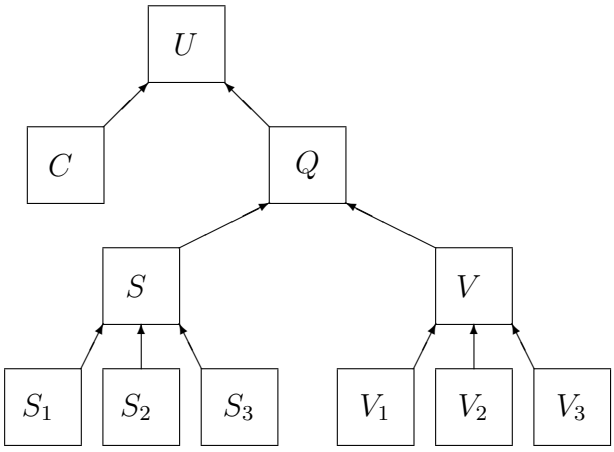

Figure 1: Utility hierarchy for the course design example.

  S_1 short term learning,
  S_2 longer-term learning,
  S_3 satisfaction,

- For the university and staff:

  V_1 staff satisfaction,
  V_2 institutional benefits,
  V_3 staff development,
  C financial cost.

As for many decision problems, the attributes of interest are in very different units and it may be difficult to establish precise trade-offs between the attributes in order to rank the various teaching choices.

2.3 Example: Utility hierarchy

The utility hierarchy is shown in Figure 1.

The overall utility node U is a binary node, combining the utility U_C for cost and the utility U_Q for quality. So the overall utility is

$$U = a_{UQ}U_Q + a_{UC}U_C + h_U U_Q U_C.$$

The "module quality" utility U_Q is formed at a binary node and is given by

$$U_Q = a_{QS}U_S + a_{QV}U_V + h_Q U_S U_V,$$

where U_S and U_V are the utilities for "Students" and "University". Each of these is an additive node which depends on three marginal utilities:

$$\begin{aligned} U_S &= a_{S1}U_{S1} + a_{S2}U_{S2} + a_{S3}U_{S3}, \\ U_V &= a_{V1}U_{V1} + a_{V2}U_{V2} + a_{V3}U_{V3}. \end{aligned}$$

	Vertex			
	$\underline{\phi}_{S1}$	$\underline{\phi}_{S2}$	$\underline{\phi}_{S3}$	$\underline{\phi}_{S0}$
a_{S1}	0.2	0.2	0.5	0.3
a_{S2}	0.4	0.7	0.4	0.5
a_{S3}	0.4	0.1	0.1	0.2

Node S (Students).

	Vertex			
	$\underline{\phi}_{V1}$	$\underline{\phi}_{V2}$	$\underline{\phi}_{V3}$	$\underline{\phi}_{V0}$
a_{V1}	0.05	0.05	0.20	0.10
a_{V2}	0.50	0.75	0.55	0.60
a_{V3}	0.45	0.20	0.25	0.30

Node V (University).

	Vertex				
	$\underline{\phi}_{Q1}$	$\underline{\phi}_{Q2}$	$\underline{\phi}_{Q3}$	$\underline{\phi}_{Q4}$	$\underline{\phi}_{Q0}$
a_S	0.890	0.500	0.890	0.500	0.695
a_V	0.110	0.500	0.305	0.305	0.305
h_Q	0.000	0.000	-0.196	0.195	0.000

Node Q (Module Quality).

	Vertex				
	$\underline{\phi}_{U1}$	$\underline{\phi}_{U2}$	$\underline{\phi}_{U3}$	$\underline{\phi}_{U4}$	$\underline{\phi}_{U0}$
a_C	0.7	0.5	0.7	0.5	0.6
a_Q	0.3	0.5	0.4	0.4	0.4
h_U	0.0	0.0	-0.1	0.1	0.0

Node U (Overall Utility).

Table 1: Trade-off parameter values.

The marginal utilities U_{S1}, U_{S2}, U_{S3}, U_{V1}, U_{V2}, U_{V3} are associated with the attributes S_1, S_2, S_3, V_1, V_2, V_3. Details of the evaluation of expected marginal utilities are given in [4].

The utility function is fully specified when we assign values to all of the trade-off parameters in the above relations. In this paper, we shall consider how to analyse the problem as an SIIH, when we are unwilling to give precise values to these trade-offs.

2.4 Using Imprecise Trade-off Parameters

One of the most difficult tasks in specifying a mutually utility independent structure is the quantification of the various trade-off parameters in the forms (2), (3) and (1), as this typically requires the comparison of intrinsically different types of costs and benefits. Therefore, it is of fundamental interest to consider problems where we are unwilling to fix on particular trade-off values or where a group of individuals must make a joint decision, and there is broad agreement on the marginal utilities, but different members of the group have different priorities when trading risks.

Although we are unwilling to place strict values on the trade-offs, there will be certain combinations of outcomes over which we are prepared to state preferences and these comparisons establish the region of the space of trade-off parameters which we must consider. We choose to elicit our imprecision in the values of the trade-off parameters $\underline{\theta}$ based on our stated preferences over utility combinations for outcomes, as this is usually more meaningful than considering directly the imprecision in the elements of $\underline{\theta}$. So, for each child node, we make a collection of pairwise comparisons between vectors of values of parent utilities (or, equivalently, the corresponding vectors of attribute values). Details are given in [3].

Some authors also consider imprecision in the marginal utility functions. Recent examples include [7] who describe a decision support system in which the imprecise multi-attribute utility function is additive and [5], who allow a multiplicative function in which a range for the value of k in (1) is determined by considering the values implied by ranges given for $a_1, \ldots, a_s$. In both cases ranges for the trade-off parameters are combined to form a rectangular space. In this paper we only consider imprecision in trade-offs and assume that the necessary expectations of marginal utilities, and in some cases their products, can be agreed. However we do not impose an arbitrary probability distribution over ranges of imprecision, or over attributes, nor do we assume a rectangular shape for the space of trade-off parameters allowed by the imprecise specification resulting from a careful elicitation process.

For each additive or binary child node, we state whichever preferences we wish between pairs of utility vectors for the parent nodes. Each stated preference places a linear constraint on the allowable choices for the trade-off parameters $\underline{\phi}_i$. We term the collection, R, of all sets of trade-off parameters consistent with each of the stated preferences the *feasible* region of choices for the trade-off parameters. In [3] we showed that the shape of the region of trade-off parameters resulting from the above elicitation scheme for an SIIH is as follows. At each additive or binary node n, we obtain a convex polyhedron R_n for the allowable values of $\underline{\phi}_n$. The regions $R_1, \ldots, R_N$ together define a

d_1	1,	3,	1,	1,	3,	2
d_2	2,	3,	1,	3,	3,	2
d_3	1,	3,	1,	3,	3,	2
d_4	1,	3,	1,	1,	1,	3
d_5	1,	3,	1,	1,	3,	3
d_6	2,	3,	1,	1,	3,	2

Table 2: Alternatives for comparison.

region R in the combined space of parameters $\underline{\theta}$, where $\underline{\theta} \in R$ if and only if $\underline{\phi}_n \in R_n$ for $n = 1, \ldots, N$. The vertices of R_n are denoted $\underline{\phi}_n^{(1)}, \ldots, \underline{\phi}_n^{(r_n)}$ and those of R are denoted $\underline{\theta}^{(1)}, \ldots, \underline{\theta}^{(r)}$. Let P be the set of vertices of R and P_n be the set of vertices of R_n.

We explained in [3] that, in the case of an IIH containing multiplicative nodes where the utilities are combined using (1) and (4), we must modify the elicitation procedure. We also described the shape of the resulting feasible set. If we are willing to choose a fixed value for k then, at each multiplicative node n, we obtain a bounded rectangular region $R_n(k)$, with vertices $\underline{\phi}_n^{(1)}, \ldots, \underline{\phi}_n^{(r_n)}$, for the remaining elements of $\underline{\phi}_n$. The shape is somewhat more complicated if the value of k is also treated as imprecise.

2.5 Example: Imprecise trade-offs

The specification of imprecise utility trade-offs in this example was described in more detail in [4]. Table 1 gives the vertex set P_i for the feasible region R_i, at each node i. For each node, a central value $\underline{\phi}_{i0}$ is also listed, which is the average of the values at each vertex.

In [4] we found that there were 50 Pareto optimal choices in this example. Of these, 37 could be eliminated because they were equivalent to other choices which were retained. We chose the value $\varepsilon = 0.012$ and, by applying our ideas of almost-preference with this value of ε, reduced the list to the six alternatives listed in Table 2. These are ordered according to our ε-preference procedure, d_6 being eliminated before d_5 and so on. The last remaining choice is d_1.

3 Boundary linear utility

3.1 Definitions and motivation

The feasible region for the trade-off parameters in a SIIH is the convex hull of a finite collection of trade-off parameters $\underline{\theta}^{(i)} \in P, i = 1, \ldots, r$. We now need a way to compare non-dominated choices over this region. Let U_i be the utility function determined by the choice of trade-offs $\underline{\theta}^{(i)} \in P, i = 1, \ldots, r$. Any function of the form

$$\bar{U}_\lambda = \sum_{i=1}^{r} \lambda_i U_i \tag{5}$$

where $\lambda = (\lambda_1, \ldots, \lambda_r)$ are non-negative constants such that $\sum_{i=1}^r \lambda_i = 1$ is termed a *boundary linear utility*. For any such $\bar{U}_\lambda$, we may identify the rule which maximises $\bar{U}_{d,\lambda} = \sum_{i=1}^r \lambda_i U_{d,i}$, where $U_{d,i}$ is the utility of alternative d with trade-off $\underline{\theta}^{(i)}$.

In [3] the boundary linear form is motivated by various axiomatic and natural requirements for the combination of group preferences. In addition to such theoretical support, the boundary linear form is easy to interpret, gives a clear comparison between different choices and leads to tractable procedures even for large numbers of alternative decisions. The choice of the λ weights can be used to emphasise or de-emphasise the importance of a particular attribute by putting more or less weight on vertices corresponding to different values for a particular trade-off.

While the set of λ weights is formally equivalent to a probability distribution over the points in P, our interpretation of the λ weights is not probabilistic but is in terms of the properties of the boundary linear utility described below and as a means for exploring the robustness of alternatives. A probability distribution over possible sets of attribute weights is used in [1] as a means of exploring sensitivity. In [2] a weight specification, known as a second order belief specification, over the ranges of imprecisely specified probabilities and expected utilities in a decision tree is used to help make a unique choice of alternative.

3.2 Properties of the boundary linear utility

Let us consider first the case of a SIIH.

There is a natural relation between Pareto optimality and Bayes rules for boundary linear utilities. In [3] we showed that, for a SIIH, a decision which is either (i) a unique Bayes decision for some $\bar{U}_\lambda$, or (ii) a Bayes decision for some $\bar{U}_\lambda$ with $\lambda_i > 0$ for $i = 1, \ldots, r$, is Pareto optimal over R.

Each weight λ_i corresponds to a complete parameter specification $\underline{\theta}^{(i)}$. It is useful to be able to relate this to weights applied to parameter specifications at individual nodes. Denote by $\lambda(i_1, \ldots, i_N)$ the weight applied to the combination of vertices $\underline{\phi}_1^{(i_1)}, \ldots, \underline{\phi}_N^{(i_N)}$ at nodes $1, \ldots, N$ respectively. Denote by $\lambda_{n,i}$ the weight applied to vertex $\underline{\phi}_n^{(i)}$ at node n. If we require that the weights applied to vertices at node n should not change if we combine this vertex with a different vertex at another node then we require

$$\frac{\lambda(i_1,\ldots,i_n,\ldots,i_N)}{\lambda(i_1,\ldots,i'_n,\ldots,i_N)} = \frac{\lambda_{n,i_n}}{\lambda_{n,i'_n}}$$

for two different vertices i_n and i'_n at node n, with $\lambda_{n,i'_n} \neq 0$. It follows that $\lambda(i_1,\ldots,i_N) = \prod_{n=1}^{N} \lambda_{n,i_n}$. Such a weight specification is called a *multiplicative weighting*. For such a specification, we may vary the weights at each node separately.

It is often helpful to equate the boundary linear form with the utility at interior trade-off values. It follows directly from the fact that R_i is a convex polyhedron that, for any $\underline{\theta}$ in R, there exists a multiplicative weighting λ such that $\underline{\theta} = \bar{\underline{\theta}}_\lambda$ and, for any multiplicative weighting λ, there exists a $\underline{\theta}$ in R such that $\underline{\theta} = \bar{\underline{\theta}}_\lambda$, where $\bar{\underline{\theta}}_\lambda = \sum_j \lambda_j \underline{\theta}^{(j)}$ and the sum is taken over all of the vertices of R. In [3] we showed that, in a SIIH, if λ is a multiplicative weighting then $\bar{U}_\lambda = U(\bar{\underline{\theta}}_\lambda)$. This result establishes a correspondence between the elements of R and the multiplicative boundary linear utilities.

From this we know that, for any $\underline{\theta}$ in R, we can find $\lambda_1,\ldots,\lambda r$ such that $U(\underline{\theta}) = \sum_i \lambda_i U_i$. Values of $\underline{\theta}$ not on the boundary of R will give $\underline{\lambda}$ values satisfying $\lambda_i > 0$ for $i = 1,\ldots,r$. Rules which are Bayes for such internal $\underline{\theta}$ values will therefore be Pareto optimal over R.

For illustration of the multiplicative weighting, consider a simple example with three marginal utilities and two additive nodes where

$$U_0 = \phi_{01}U_1 + \phi_{02}U_2,\ U_1 = \phi_{13}U_3 + \phi_{14}U_4$$

and at each of the two nodes we have two alternative parameter specifications, corresponding to the vertex values. The two values for ϕ_{01} are ϕ_{011} and ϕ_{012} etc. Thus R has four vertices. Assign weight λ_{jk} to the vertex where node 0 takes parameter specification j and node 1 takes parameter specification k. The coefficient of U_3 in U_0 is now

$$\begin{aligned}\Phi_3 &= \{(\lambda_{11}+\lambda_{12})\phi_{011} + (\lambda_{21}+\lambda_{22})\phi_{012}\} \\ &\quad \times \{(\lambda_{11}+\lambda_{21})\phi_{131} + (\lambda_{12}+\lambda_{22})\phi_{132}\}.\end{aligned}$$

Now introduce weights on the parameter values at the individual nodes and calculate the overall weights from these so that $\lambda_{11} = \lambda^\star_{01}\lambda^\star_{11}, \lambda_{12} = \lambda^\star_{01}\lambda^\star_{12}, \lambda_{21} = \lambda^\star_{02}\lambda^\star_{11}, \lambda_{22} = \lambda^\star_{02}\lambda^\star_{12}$, where $\lambda^\star_{01}$ is the weight on the first parameter set at node 0 etc. and the weights at each node sum to 1. The coefficient of U_3 now simplifies to

$$\begin{aligned}\Phi_3 &= \{\lambda^\star_{01}\phi_{011} + \lambda^\star_{02}\phi_{012}\}\{\lambda^\star_{11}\phi_{131} + \lambda^\star_{12}\phi_{132}\} \\ &= \lambda^\star_{01}\lambda^\star_{11}\phi_{011}\phi_{131} + \lambda^\star_{01}\lambda^\star_{12}\phi_{011}\phi_{132} \\ &\quad +\lambda^\star_{02}\lambda^\star_{11}\phi_{012}\phi_{131} + \lambda^\star_{02}\lambda^\star_{12}\phi_{012}\phi_{132} \\ &= \lambda_{11}\phi_{011}\phi_{131} + \lambda_{12}\phi_{011}\phi_{132} + \lambda_{21}\phi_{012}\phi_{131} \\ &\quad +\lambda_{22}\phi_{012}\phi_{132}\end{aligned}$$

a weighted average of the coefficients at the four vertices, as required.

3.3 Boundary linear utility in a general IIH

The boundary linear utility is easily extended to the case where a hierarchy contains multiplicative nodes where the utilities are combined as in (1) and (4) provided that a precise value of the parameter k is used. The extension to the case where k is imprecisely specified is discussed in [3] where we showed that, in any IIH, for any $\underline{\theta}$ in R there exists a multiplicative weighting λ such that $U(\underline{\theta}) = \bar{U}_\lambda$, thus generalising the correspondence between the elements of R and the multiplicative boundary linear utilities.

3.4 Example: Boundary linear utility

With equal λ weights on all vertices, the alternative which maximises $\mathrm{E}(\bar{U}_\lambda)$ is rule d_1 which gives $\mathrm{E}(\bar{U}_\lambda) = 0.5120$. The central point $\underline{\theta}_0$, at which $U(\underline{\theta}) = \bar{U}_\lambda$, is given by the centres of each range as given in Table 1.

The λ weights could be varied to change the emphasis on different attributes. For example, at node U the coefficient of financial cost varies between 0.5 and 0.7. Putting more weight on all vertices where the coefficient was 0.7 would emphasise this attribute, whereas more on all vertices where it was 0.5 would de-emphasise it. For illustration we changed the weights to 2:1 in favour of 0.7 and 2:1 in favour of 0.5. In each case rule d_1 maximised $\mathrm{E}(\bar{U}_\lambda)$ giving values of 0.5096 and 0.5144 respectively. This increases our confidence in the choice of d_1.

Sometimes, we may uniquely choose a collection of λ weights under the guidance of one of the formal arguments in [3]. However, usually we will want to consider the robustness of our choice to variation in λ, which we now address more formally.

4 Exploring sensitivity

4.1 General comments

The boundary linear utility gives us an approach to choosing between alternative rules. However, while

any given boundary linear utility function identifies a "best" alternative, we would usually prefer an alternative which is robust in the sense that it behaves well compared to most alternatives over most of the range of trade-off parameters. We now consider how such robustness may be assessed.

When we have chosen a multiplicative boundary linear utility $\bar{U}_\lambda = \sum_{i=1}^{r} \lambda_i U_i$, we find the decision $d^\star$ which maximises expected utility, under $\bar{U}_\lambda$. We also define a 'central' parameter specification $\underline{\theta}_0 = \sum_{i=1}^{r} \lambda_j \underline{\theta}^{(j)}$ where this sum is taken over the elements of P. From Section 3.2 we know that, in a SIIH, when λ is a multiplicative weighting, $U(\underline{\theta}_0) = \bar{U}_\lambda$. Thus, we can explore sensitivity in two ways. First, we can see how much we must change the λ weightings in order to alter our choice of best decision and secondly, at least in a SIIH, we can see how far we must move away from the central value θ_0, to alter our choice. Effectively, this establishes two separate but related sensitivity metrics. The former is concerned solely with the relative importance of the various vertices of the trade-off space, irrespective of their Euclidean values, while the latter reflects the actual Euclidean distances between alternative trade-off parameters.

The investigations described below are designed to assess the robustness of our decision to the choice of trade-off. At each step, if the analysis suggests that there are other alternatives which perform substantially better than our selected rule over much of the trade-off space, then we may repeat the steps, substituting the suggested alternatives, to see whether a more robust choice of rule may be found.

Suppose, in what follows, that we have a set D of alternatives for comparison with $d^\star$. This set may be a subset of the Pareto optimal choices formed using the methods in [4]. Suppose also that we have chosen a small increment $\varepsilon > 0$ which we tolerate in comparing utilities, as discussed in [4].

4.2 Volume sensitivity

A first general robustness measure is as follows. For each alternative in D, we compute the volume of λ-space, as a proportion of the total volume within which $\sum \lambda_j = 1$, over which the difference in utility between that alternative and $d^\star$ is at least ε. If this proportion is very small, then this suggests that $d^\star$ is robust against that alternative.

Having assessed global sensitivity over the whole hierarchy, we may repeat the analysis in any subhierarchy. For any child node i, with utility U_i, we may find the proportion of the permissible λ-space for the vertices of the feasible region of parameters in the sub-hierarchy under i in which the difference in expectations of U_i between an alternative and $d^\star$ is at least ε.

To do these analyses we need to be able to compute the volume of λ-space which satisfies a condition

$$g(d_1, d_2) = \bar{U}_\lambda(d_1) - \bar{U}_\lambda(d_2) > x \qquad (6)$$

for some specified x, where d_1 and d_2 are two choices. Let $\underline{d} = (d_1, \ldots, d_e)$ and $\underline{\bar{U}}_\lambda^{(n)}(\underline{d}) = (\bar{U}_\lambda^{(n)}(d_1), \ldots, \bar{U}_\lambda^{(n)}(d_e))$, where $\bar{U}_\lambda^{(n)}(d_j)$ is the boundary linear utility evaluated at node n with weights $\underline{\lambda}$ over the subhierarchy $H(n)$ under n. To evaluate the volume satisfying (6), we can make use of the following analogy.

If we gave $\underline{\lambda}$ a uniform distribution over its feasible region then the required volume would be the probability that (6) is satisfied. The utility hierarchy can then be interpreted as a graph in which the probability distribution of the utility difference between any two decisions at a child node, given the values of the parent utilities, would depend only on the distribution of the tradeoff parameters at the child node. Thus we can evaluate the distribution of $\underline{\bar{U}}_\lambda^{(n)}(\underline{d})$ higher in the hierarchy through a chain of conditional distributions. See, e.g., [9].

Specifically, the density of $\underline{\bar{U}}_\lambda^{(n)}(\underline{d})$, the values at a child node n with parents $n_1, \ldots, n_s$, is

$$f_n(\underline{\bar{U}}_\lambda^{(n)}(\underline{d})) = \int \cdots \int \left\{ f_{n|H(n)}[\underline{\bar{U}}_\lambda^{(n)}(\underline{d}) \mid \underline{\bar{U}}_\lambda^{(n^*)}(\underline{d})] \prod_{i=1}^{s} f_{n_i}(\underline{\bar{U}}_\lambda^{(n_i)}(\underline{d}) \right\} d\underline{\bar{U}}_\lambda^{(n^*)}(\underline{d}) \qquad (7)$$

and

$$\Pr(g_n > x) = \int \cdots \int \left\{ \Pr[g_n > x \mid \underline{\bar{U}}_\lambda^{(n^*)}(\underline{d})] \prod_{i=1}^{s} f_{n_i}[\underline{\bar{U}}_\lambda^{(n_i)}(\underline{d})] \right\} d\underline{\bar{U}}_\lambda^{(n^*)}(\underline{d}) \qquad (8)$$

where $\underline{\bar{U}}_\lambda^{(n^*)}(\underline{d}) = \underline{\bar{U}}_\lambda^{(n_1)}(\underline{d}), \ldots, \underline{\bar{U}}_\lambda^{(n_s)}(\underline{d})$ and $f_{n|H(n)}[\underline{\bar{U}}_\lambda^{(n)}(\underline{d}) \mid \underline{\bar{U}}_\lambda^{(n^*)}(\underline{d})]$ is the conditional density given the values of the boundary linear utilities evaluated at the parent nodes for the elements of $\underline{d}$.

Starting with the children of the marginal nodes, the distribution of $\underline{\bar{U}}_\lambda^{(n)}(\underline{d})$ is evaluated node-by-node up

the hierarchy using (7). At a child node n with r_n vertices we have, from (5) and (6),

$$g_n(d_1, d_2) = \sum_{i=1}^{r_n} \lambda_{n,i}[U_i^{(n)}(d_1) - U_i^{(n)}(d_2)]$$

where $U_i^{(n)}(d)$ is evaluated at vertex i of node n. Thus $g_n(d_1, d_2) = x$ defines a plane in λ-space which may cut the feasible region. The conditional probability $\Pr[g_n > x \mid \underline{\bar{U}}_\lambda^{(n^*)}(\underline{d})]$ in (8) is then a proportion of the volume of the feasible polyhedron which can be determined by finding where $g_n = x$ cuts the edges.

Similarly, the conditional probability $\Pr[\bar{U}_\lambda^{(n)}(d_1) < x \mid \underline{\bar{U}}_\lambda^{(n^*)}(\underline{d})]$ is the proportion of the volume of the feasible polyhedron at node n cut off by $\bar{U}_\lambda^{(n)}(d_1) = x$, with the parent utilities fixed. Differentiating this probability with respect to x gives the conditional density of $\bar{U}_\lambda^{(n)}(d_1)$. Then fixing $\bar{U}_\lambda^{(n)}(d_1) = x_1$ imposes a linear constraint on $\lambda_{n,1}, \ldots, \lambda_{n,r_n}$ and reduces the dimension of the feasible region by 1. By considering the intersection of $\bar{U}_\lambda^{(n)}(d_2) = x_2$ with this reduced region we can find the conditional distribution of $\bar{U}_\lambda^{(n)}(d_2)$ given $\bar{U}_\lambda^{(n)}(d_1) = x_1$. If required, we can continue this process for $d_3, \ldots, d_{r_n-1}$. For $j > r_n - 1$, $\bar{U}_\lambda^{(n)}(d_j)$ is then a deterministic function of $\underline{\bar{U}}_\lambda^{(n)}(d_1, \ldots, d_{r_n-1})$. In this way we can find the conditional density $f_{n|H(n)}[\underline{\bar{U}}_\lambda^{(n)}(\underline{d}) \mid \underline{\bar{U}}_\lambda^{(n^*)}(\underline{d})]$ in (7).

4.3 Example: Volume sensitivity

We have identified the choice d_1 under the utility with equal weightings at each vertex in P. We now consider the sensitivity of that choice, following the steps in Section 4.

We computed the volume of λ-space, as a proportion of the total volume within which $\sum \lambda_j = 1$, over which the difference in utility between alternative d_1 and each of the other retained alternatives is at least $-\varepsilon$, at our chosen value of 0.012. We concluded that the volume over which the difference in favour of any alternative over d_1 is greater than ε is less than 0.01% of the total volume and therefore that d_1 is a robust choice. (The proportion is nonzero, since we know that the difference is greater than ε at some of the vertices. However the region of λ space which we are exploring is a simplex of very high dimension and the neighbourhoods of the vertices of this simplex contribute only a tiny fraction of the total volume.)

Next we computed the proportions of λ-volume over which each alternative's boundary linear utility exceeded that of d_1 by at least ε for each of the non-marginal nodes in the hierarchy. Table 3 gives the results. The results show that the challenge to d_1 seems to be based in node V. The apparent main challengers, rules d_2 and d_6, differ only in Unit 4 which is given by lectures in d_6 and open learning in d_2. According to the elicited expectations, d_6 thus favours the students more.

		Node			
		U	Q	S	V
Rule	d_2	0.000	1.000	0.103	1.000
	d_3	0.000	0.001	0.000	1.000
	d_4	0.000	0.000	0.000	0.000
	d_5	0.000	0.000	0.000	0.000
	d_6	0.000	1.000	1.000	1.000

Table 3: Proportions of λ volume where the utility difference is at least ε at non-marginal nodes.

	Choice					
	d_1	d_2	d_3	d_4	d_5	d_6
C	0.484	0.416	0.476	0.544	0.536	0.424
S_1	0.497	0.447	0.463	0.433	0.400	0.480
S_2	0.578	0.577	0.528	0.420	0.370	0.627
S_3	0.800	0.900	0.800	0.800	0.800	0.900
V_1	0.533	0.467	0.433	0.500	0.400	0.567
V_2	0.433	0.667	0.533	0.300	0.400	0.567
V_3	0.467	0.717	0.583	0.333	0.450	0.600

Table 4: Values of expected utilities at the marginal nodes.

4.4 Distances in λ-space

Next, for each alternative in D, we identify those vertices where the difference in utility between that choice and $d^\star$ is at least ε. For each of these vertices, we find the distance, in λ-space, in the direction of the vertex, between $\underline{\lambda}_0$, our original λ specification, and the point where the difference in boundary linear utility between that choice and $d^\star$ first reaches ε. Let $\parallel \underline{\lambda} \parallel = \sqrt{\underline{\lambda}'\underline{\lambda}}$, where $\underline{\lambda}'$ is the transpose of $\underline{\lambda}$. We find $t \parallel \underline{\lambda}_v - \underline{\lambda}_0 \parallel$, where $\underline{\lambda}_v$ is the λ vector for a vertex, $t = \{\delta(\underline{\lambda}_0) + \varepsilon\}/\{\delta(\underline{\lambda}_0) - \delta(\underline{\lambda}_v)\}$ and $\delta(\underline{\lambda})$ is the difference in boundary linear utility at $\underline{\lambda}$. Large values of these distances suggest robustness of $d^\star$. In this metric, the distance between any two vertices is $\sqrt{2}$.

4.5 Example: Distances in λ-space

Table 4 shows the values of the expected marginal-node utilities for the members of D and Table 6 shows at which marginal node each alternative is superior to $d^\star$. Table 5 lists the vertices where the difference in utility between one of the other alternatives and

$d^\star$ is at least ε. The vertices are numbered for easy reference. The vertices can be identified using the numbering of the vertices at each node, which is the same as in Table 1. Table 5 then gives the distances, from the original $\underline{\lambda}$ specification towards these vertices, to reach points where the difference in utility between one of the other alternatives and $d^\star$ is at least ε. Most of the distances are large. There are a few exceptions, notably for rule d_2 at vertices 46 and 47. Rule d_2 is the retained option with the least dependence on traditional lectures and at these vertices relatively little weight is placed on financial cost but relatively great weight is placed on institutional benefit. To put the distances in context, observe that each original λ value is approximately 0.007. The move required for vertex 46 changes λ_{46} to approximately 0.6 and therefore the average of the other λ values is less than 0.003 or 0.5% of λ_{46}. There seems to be little reason here to change our conclusion that d_1 is a robust choice. Notice also how the pattern of marginal nodes in common between rules in Table 6 tends to be repeated with vertices in common in Table 5.

4.6 Sensitivity in the θ-metric

We can quantify sensitivity in the θ-metric by looking at the effect of general movement away from $\underline{\theta}_0$ as follows. Let the elements of P be $\underline{\theta}^{(1)}, \ldots, \underline{\theta}^{(r)}$. Define the *scaled range* R_t to be the convex hull of P_t, the elements of which are given by $\underline{\theta}_t^{(i)} = \underline{\theta}_0 + t(\underline{\theta}^{(i)} - \underline{\theta}_0)$ for $t \geq 0$. We may think of this as expanding a volume (in the θ-metric) centred on $\underline{\theta}_0$ until a boundary of the region of optimality of $d^\star$ is reached. An obvious extension of Lemma 2 in [3] shows that this boundary will be reached first at an element of P_t so we only need to make comparisons at the vertices. For each element of D we evaluate, at each of a range of values of t up to 1, the maximum over P_t of the difference in expected utility compared with $d^\star$ and plot these values against t. This plot will serve as an indication of over how large a range around $\underline{\theta}_0$ we can judge $d^\star$ to be robust. This approach may be compared with that of [8] in which the sensitivity of a preferred alternative is measured using the minimum distance (in some metric) to a point in the parameter space at which another alternative becomes preferable.

4.7 Example: Sensitivity in the θ-metric

Figure 2, shows one of the range expansion plots. The horizontal axis is the expansion factor t. The vertical axis is the difference in expected utility between an alternative, in this case d_2, and $d^\star$, in this case d_1. At each value of the expansion factor the values at the 144 vertices of the range were calculated and the maximum, minimum, median and upper and lower quartiles of these 144 values are plotted.

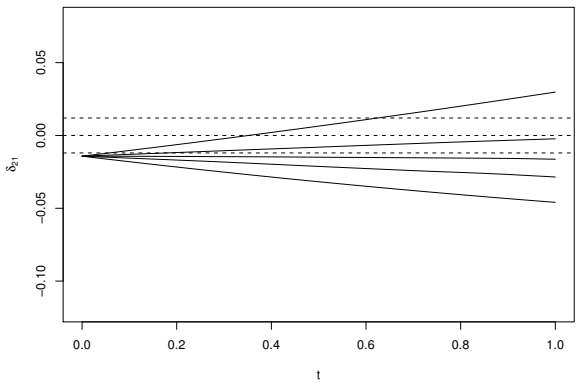

Figure 2: Expansion with respect to all parameters. Maximum, quartiles and minimum of the difference in expected utility between d_2 and d_1 at 144 vertices, against expansion factor t. Reference lines are given at zero and $\pm\varepsilon$.

Choice	Marginal Node					
d_2			S_3		V_2	V_3
d_3					V_2	V_3
d_4	C					
d_5	C					
d_6		S_2	S_3	V_1	V_2	V_3

Table 6: Marginal nodes at which alternatives are superior to d_1.

From Figure 2, we see that d_2 does substantially worse than d_1 over most of the range but possibly better for large t. Similar plots for the other alternatives show that none of the other rules does much better than d_1 over any of the range and some do much worse in some of the range. Generally the maximum difference only exceeds ε towards the end of the range. We conclude that d_2 is the only alternative to d_1 worth further consideration.

5 Conclusion

In [3], [4] and this paper we have described an approach to multi-attribute decision analysis where the trade-offs between attributes are not precisely specified. Imposing the condition of utility independence makes the dimensionality of the trade-off specification finite and allows us to work in terms of ranges for trade-off parameters. However, by imposing this condition only at the nodes of a utility hierarchy we can relax the requirement for mutual utility indepen-

	Vertex for Node				Distance for Alternative			Vertex for Node				Distance for Alternative		
Vertex	U	Q	S	V	d_4	d_5	Vertex	U	Q	S	V	d_2	d_3	d_6
1	1	1	1	1	0.591		46	2	2	1	1	0.354	0.815	0.710
2	1	1	1	2	0.591		47	2	2	1	2	0.371	0.919	0.710
3	1	1	1	3	0.567		48	2	2	1	3	0.672		0.967
7	1	1	3	1	0.954		49	2	2	2	1	0.518		0.978
8	1	1	3	2	0.954		50	2	2	2	2	0.548		0.978
9	1	1	3	3	0.906		52	2	2	3	1	0.638		
10	1	2	1	1		0.698	53	2	2	3	2	0.681		
11	1	2	1	2		0.732	64	2	4	1	1	0.468		0.779
12	1	2	1	3	0.843	0.867	65	2	4	1	2	0.492		0.782
16	1	2	3	1		0.994	66	2	4	1	3	0.892		
19	1	3	1	1	0.654		67	2	4	2	1	0.929		
20	1	3	1	2	0.656		68	2	4	2	2	0.992		
21	1	3	1	3	0.605		118	4	2	1	1	0.637		
28	1	4	1	1	0.942	0.809	119	4	2	1	2	0.673		
29	1	4	1	2	0.938	0.840	121	4	2	2	1	0.991		
30	1	4	1	3	0.782	0.983	136	4	4	1	1	0.878		
							137	4	4	1	2	0.930		

Table 5: Distances to points where the utility difference is at least ε.

dence between all attributes. In our earlier papers we discussed how to reduce the number of alternatives for consideration and how to make a robust choice. In this paper we have considered the examination of sensitivity of our choice, in particular using the boundary linear utility.

The example illustrated the use of our methods. We gained a better understanding of the issues which are important in making our choice and greater confidence in our selection of d_1. We saw that d_2 posed the most important challenge to the choice of d_1 and identified node V as the main basis for this challenge.

We believe that, in many difficult decision problems where a range of trade-off specifications must be considered, our methods could lead to the selection of a choice which is, in practical terms, close to optimal everywhere in the range.

References

[1] J. Butler, J. Jia and J. Dyer. Simulation techniques for the sensitivity analysis of multi-criteria decision models. *European Journal of Operational Research*, 103:531–546, 1997.

[2] M. Danielson, L. Ekenberg and A. Larsson. Distribution of expected utility in decision trees. *International Journal of Approximate Reasoning*, 46:387–407, 2007.

[3] M. Farrow and M. Goldstein. Trade-off sensitive experimental design: a multicriterion, decision theoretic, Bayes linear approach. *Journal of Statistical Planning and Inference*, 136:498–526, 2006.

[4] M. Farrow and M. Goldstein. Almost-Pareto decision sets in imprecise utility hierarchies. *Journal of Statistical Theory and Practice*, 3:137–155, 2009.

[5] A. Jiménez, S. Ríos-Insua and A. Mateos. A decision support system for multiattribute utility evaluation based on imprecise assignments. *Decision Support Systems*, 36:65–79, 2003.

[6] R. L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Trade-offs*, New York: John Wiley & Sons, 1976.

[7] A. Mateos, A. Jiménez and S. Ríos-Insua. Modelling individual and global comparisons for multiattribute preferences. *Journal of Multi-Criteria Decision Analysis*, 12:177–190, 2003.

[8] D. Ríos-Insua and S. French. A framework for sensitivity analysis in discrete multi-objective decision-making. *European Journal of Operational Research*, 54:176–190, 1991.

[9] J. Q. Smith. Statistical principles on graphs. In *Influence Diagrams, Belief Nets and Decision Analysis*, R. M. Oliver and J. Q. Smith (eds), New York: John Wiley & Sons, 89–120, 1990.