Department for Environment, Food & Rural Affairs

Use of Multi-Criteria Analysis in Air Quality Policy

A Report Prepared for the Department for Environment, Food and Rural Affairs by:

Catalyze

Larry Phillips
Adrian Stock

5 November 2003

Contents

1. Executive Summary

This project was jointly funded by DEFRA and the Environment Agency to explore the use of multi-criteria decision analysis (MCDA) to inform the UK Government's air quality strategy (AQS) objectives. Specifically, these objectives were to be met:

1. To assess how MCDA techniques compare with currently-employed CBA techniques employed by DEFRA.
2. To demonstrate the benefits of using MCDA techniques as a supplement to existing processes.
3. To identify areas of AQS work where MCDA techniques can most appropriately be used.

Two studies explored the use of MCDA. Both studies involved workshops attended by specialists who could contribute their expertise in assessing the various benefits of the options being considered. The workshops were facilitated by the authors of this report, and interactive computer-based MCDA modelling was employed during the workshops.

The first study concentrated on how portfolios of options could be constructed to obtain best value-for-money. Two separate one-day workshops were held. At the first workshop, participants created a 'top-down' model enabling a portfolio of best policies to be constructed for three geographical areas of the UK. At the second workshop, the group developed a 'bottom-up' model showing how a portfolio of abatement measures required of polluters could be designed. In particular, the model established the value-for-money priority of five options.

The second study used MCDA to appraise five policy options which varied in the severity of target performance for SO_2 concentration and compliance in their effect on Sites of Special Scientific Interest. Available information provided costs of abatements for a range of industry sectors. Participants in the workshop appraised the options on four criteria, SSSI At Risk (extent that damage to plants is reduced), Health Effects of Air Pollution (number of deaths and hospital admissions avoided), Non-Designated SSSI Benefits (secondary benefits to flora and fauna) and Building Damages (reduction in damage to buildings). The final model demonstrated how MCDA handles a mix of financial and non-financial objectives, creating an overall preference ordering of the options, establishing an efficient frontier consisting of the best options for a given cost, and enabling extensive sensitivity analyses to be carried out to see the effects on the overall ordering of imprecision in the input data and differences of opinion about judgements of value. In addition, it showed how to establish the key advantages and disadvantages of selected options, and how to find the important ways in which pairs of options differ.

Sensitivity analyses conducted on the models showed that considerable variations in the input data, either because of imprecision in the underlying data or disagreements between participants, did not affect the overall results. Although the models were incomplete, they still displayed considerable robustness to changes in the inputs.

The study extracted 20 findings relevant to the project's objectives. In summary, they are:

1. Different uses of language between MCDA and CBA can cause considerable confusion.
2. MCDA can capture any set of criteria, monetary and non-monetary.
3. CBA can provide the monetary inputs to MCDA; they are not alternative approaches.
4. MCDA combines social and technical processes.
5. MCDA provides an analytical structure for comparing monetary and non-monetary outputs.
6. In MCDA, human judgement is required to establish relative weights of the criteria.
7. MCDA graphs, which are typical outputs, aid understanding.
8. Precision in weights is not required in MCDA.
9. MCDA provides methods for discovering the key advantages and disadvantages of an option, and the important ways it differs from other options.
10. MCDA's focus on group modelling can lead to different results from CBA conducted in the 'back room'.
11. Proper scaling techniques are required in MCDA.

12. MCDA modelling workshops require a supportive physical environment.
13. The MCDA portfolio model can improve the efficiency of budget allocation.
14. Debates about what criteria to include in an MCDA can be informed by desirable properties discussed in the MCDA literature.
15. The distinction between inputs and outputs in the MCDA portfolio model helps to clarify what are costs and what are benefits.
16. MCDA modelling can accommodate 'top-down' portfolios of policies and 'bottom-up' portfolios of concrete courses of action, here, 'measures'.
17. Weighting in portfolio MCDA is more complex than in simple appraisal MCDA.
18. Displaying the Δbenefit/Δcost triangles for each option makes clear the extent of extra benefit expected from an extra investment.
19. Portfolio MCDA modelling overcomes the inefficient use of resources that results from 'silo' decisions.
20. Sensitivity analysis plays a crucial role in MCDA, as it does in CBA.

These findings suggest that CBA and MCDA techniques are different, but complementary: MCDA can extend CBA by effectively incorporating criteria that are difficult or impossible to monetise using CBA techniques. As applied to air quality policy options, MCDA can be used for appraisal against any set of important criteria, for constructing portfolios 'top-down' of policy options across different areas (geographical, functional, or any other category), and for constructing portfolios of 'bottom-up' abatement measures.

The best way to explore further the contribution of MCDA to the AQS would be to involve MCDA specialists as part of a team dealing with the current realities of government policy and Environment Agency measures. Application of MCDA would likely require a combination of back-room modelling and facilitated workshops.

2. Background

In October 2002, the Department for Environment, Food and Rural Affairs (DEFRA) and the Environment Agency (EA) contacted Catalyze to learn more about Multi-Criteria Decision Analysis (MCDA) and how it might inform the Air Quality Strategy (AQS) work. This was prompted in part by the difficulty in cost benefit analysis (CBA) of placing a monetary value on the health and non-health benefits of reductions in air pollution, an issue raised by the National Audit Office which led them to recommend considering the use of multi criteria analysis to help in the setting of air quality standards[1].

Over a series of meetings, a pilot study was designed based on two policy areas within the AQS to explore the application of MCDA techniques during the appraisal of policy options. Specifically, the objectives are:

1. To assess how MCDA techniques compare with currently-employed CBA techniques employed by DEFRA.
2. To demonstrate the benefits of using MCDA techniques as a supplement to existing processes.
3. To identify areas of AQS work where MCDA techniques can most appropriately be used.

Two studies were carried out in three workshops to provide material for meeting the objectives.

This report is the final summary of the pilot study.

2.1 Key Players

Many members of the Interdepartmental Group on Costs and Benefits (IGCB(AQ)) and others have given time and insight to the pilot study. A core team drawn from DEFRA, EA and the IGCB group guided all aspects of the project:

2.2 Report Scope

This report incorporates the totality of work on the pilot study. Summaries of the three models comprising the two studies are available separately.

Section 3 provides a summary of both MCDA and CBA, and compares the two approaches. The method used in the pilot study is explained in Section 4. Section 5 presents the ecosystems appraisal model, while Section 6 gives the work on the air quality portfolio models. Both sections include findings relevant to the three objectives of the study. Section 7 gives our conclusions about how MCDA can be used, particularly to inform policy decision making about the AQS. A quick overview of our findings can be obtained by reading the 20 shaded boxes.

[1] "Policy Development: Improving Air Quality", National Audit Office, October 2001.

3. MCDA and CBA compared

In this section, we compare MCDA with CBA. Two documents form the basis for the comparison. MCDA, as one form of multi-criteria analysis (MCA), is outlined in Chapter 6 of *Multi-Criteria Analysis: A Manual*[2]. CBA is presented in *The Green Book: Appraisal and Evaluation in Central Government*[3]. Each of these documents is, of course, itself based on considerable bodies of literature, and further guidance on conducting an MCDA or CBA can be found elsewhere. For example, more specific guidance on cost benefit work is given by the Cabinet Office's Regulatory Impact Unit (see www.cabinet-office.gov.uk/regulation). It is important to recognise that we are comparing more than MCDA and CBA; we are also considering how those technical approaches are implemented. While the favoured technical approach at the heart of *The Green Book* is CBA, the document is more widely concerned with appraisal and evaluation. The same stance is taken in this report as regards MCDA.

MCDA arose from decision analysis[4] , which provides models of decisions, uncertain events and consequences, usually using decision trees as the structural representation, to help an individual make a decision. Later, the theory was extended to accommodate decisions with multiple objectives[5]. MCDA now finds its professional home in the discipline of operational research.

CBA, on the other hand, has always been associated with economics[6]. The focus is on society, using welfare economics as the justification for ensuring that whatever decision is finally taken, it provides net benefits to everyone, or if some people gain while others lose, then the best option is the one with the highest risk-adjusted sum of all costs and benefits to all parties. Of course, this assumes that benefits can somehow all be expressed in monetary terms; herein lies the justification for supplementing CBA with MCDA, for the latter does not use money as a common unit for all benefits. Instead, it uses 'value', expressed on interval scales, which typically extend from 0 to 1.0, or 0 to 100.

Curiously, although MCDA and CBA appear superficially to be very different, their origins in decision theory and economic theory, respectively, share the same foundation, namely expected utility theory[7]. This theory says that a rational decision maker considering several courses of action should anticipate the possible consequences of each action, assign utilities to represent the strength of preference for each consequence and assess probabilities for the chance of each consequence occurring. The utilities are then weighted (multiplied) by the probabilities and the products summed for each action. The decision maker should then choose the course of action with the highest expected (probability weighted) utility. From that seemingly impractical prescription, the applied technology known as decision analysis has emerged. While this theory is implicit in CBA, it is actively used in decision theory.

Proponents of both MCDA and CBA would agree on the purpose of these approaches: to appraise and evaluate proposals, projects or options. Where they differ is in how this is done. The following sections summarise the steps associated with each approach, and compare them. Later, in Section 5, the comparisons will be illustrated using the material on air quality policy obtained in the three workshops.

We assume that most readers of this report are familiar with CBA but not necessarily MCDA. To them we suggest reading Chapter 6 of *Multi-Criteria Analysis: A Manual*. For a very rapid overview, read the simple illustrative case study found in the boxes of the chapter.

---

[2] Dodgson, J., Spackman, M., Pearman, A., & Phillips, L. (2000). *Multi-Criteria Analysis: A Manual*. London: Department of the Environment, Transport and the Regions. Available at www.odpm.gov.uk/about/multicriteria/index.htm.

[3] *The Green Book: Appraisal and Evaluation in Central Government.* (2003). London: The Stationery Office.

[4] Raiffa, H. (1968). *Decision Analysis*. Reading, MA: Addison-Wesley.

[5] Keeney, R. L., & Raiffa, H. (1976). *Decisions With Multiple Objectives: Preferences and Value Tradeoffs*. New York: John Wiley.

[6] Layard, R. (Ed.). (1972). *Cost-Benefit Analysis*. London: Penguin. A more recent treatment is Boardman, A., Greenberg, D., Vining, A., & Weiner, D. (2001). *Cost-Benefit Analysis: Concepts and Practice, 2nd edition*. Upper Saddle River, NJ: Prentice Hall..

[7] Savage, L. J. (1954/1972). *The Foundations of Statistics*. New York: Wiley. Second edition in 1972, New York: Dover.

3.1 MCDA Process

The MCA Manual provides an eight-step summary of the steps characterising most MCDA projects. Decision theory provides a universal language that, we discovered in this project, is sometimes at variance with similar language used in government. To help the reader's understanding of Table 1, a brief introduction to that language may be helpful[8]. Imagine you are considering purchasing a new laptop computer. The different makes and models you are considering are *options*. Some computers meet your *objectives* of minimising cost while maximising performance better than others, but no one computer seems best in all respects. The various performance features that are relevant to your objectives are called *criteria*: purchase cost, warranty cost, weight, size of hard drive, processing speed, etc. You *value* the performance of a particular laptop on a given criterion insofar as it helps to achieve your objectives. Some criteria are relatively more important to you than others, so they will exhibit more *weight* in your final decision. Eventually you purchase a computer, and it is then that the *consequences* of your purchase are confirmed: you experience the actual performance of the machine.

Table 1. A summary of the MCDA process. Reproduced from the *MCA Manual*, p. 50.

Applying MCDA: Detailed steps
1. Establish the decision context. 1.1 Establish aims of the MCDA; identify decision makers and other key players. 1.2 Design the socio-technical system for conducting the MCDA. 1.3 Consider the context of the appraisal.
2. Identify the options to be appraised.
3. Identify objectives and criteria. 3.1 Identify criteria for assessing the consequences of each option. 3.2 Organise the criteria by clustering them under high-level and lower-level objectives in a hierarchy.
4. 'Scoring'. Assess the expected performance of each option against the criteria. Then assess the value associated with the consequences of each option for each criterion. 4.1 Describe the consequences of the options. 4.2 Score the options on the criteria. 4.3 Check the consistency of the scores on each criterion.
5. 'Weighting'. Assign weights for each of the criteria to reflect their relative importance to the decision.
6. Combine the weights and scores for each option to derive an overall value. 6.1 Calculate overall weighted scores at each level in the hierarchy. 6.2 Calculate overall weighted scores.
7. Examine the results.
8. Sensitivity analysis. 8.1 Conduct a sensitivity analysis: do other preferences or weights affect the overall ordering of the options? 8.2 Look at the advantage and disadvantages of selected options, and compare pairs of options. 8.3 Create possible new options that might be better than those originally considered. 8.4 Repeat the above steps until a 'requisite' model is obtained.

[8] The Glossary on page 55 provides complete definitions.

Following all these steps may not be necessary. For some cases, the analysis stops at step 4.1, describing the consequences after the previous steps have been completed. For example, the new approach to appraising trunk roads in England[9] stops at this point. Each proposed road project is shown in an Appraisal Summary Table that gives the appraisals on 14 criteria under the five objectives of accessibility, safety, economy, environment and integration. Integrating this information is left to the decision maker(s). We consider this to be an MCA, but not a full MCDA. An MCDA is typically conducted in an iterative fashion, with much looping back to previous steps, revising the model, gaining new insights, further modifying the model, until a requisite representation of the problematic situation is attained.

3.2 CBA Process

A similar table is not available in *The Green Book*. However, chapter 2 and the chapters themselves provided sufficient information for us to create a table of steps.

Table 2: A summary of the CBA process, extracted from *The Green Book*.

Applying CBA: Detailed steps
1. Provide a justification for action.
1.1 Establish the rationale for intervention.
1.2 Determine whether intervention is warranted.
1.3 Identify the scope of the issues involved and the basis for government action.
2. Set objectives
2.1 Establish what the proposals are intended to achieve
2.2 Create a hierarchy of outcomes, outputs and targets.
3. Appraise the options to establish a Base Case.
3.1 Prepare a list of the range of possible actions to achieve the objectives.
3.2 Value the costs and benefits of the options.
3.3 Adjust, where necessary, the costs and benefits to take account of distributional aspects and relative price changes
3.4 Discount all costs and benefits to 'present values'.
3.5 Adjust for differences in tax between options.
3.6 Adjust for risk and optimism, and consider the impact on the Base Case of alternative scenarios and changes in key variables.
3.7 Calculate the net benefits or costs.
3.8 Consider unvalued costs and benefits using 'weighting and scoring'.
4. Develop and implement the solution.
4.1 Identify the 'best' option.
4.2 Determine the affordability of options.
4.3 Through consultation, refine the 'best' option or options into a solution.
4.4 Devise implementation plans.
4.5 Present the results
4.6 Implement the solution
4.7 Track the success of the policy, programme or project in achieving its objectives.

It is common in UK Government circles to distinguish between appraisal, which is forward looking for proposed policies, programmes or projects, from evaluation, which looks backward to determine whether an existing policy, programme or project is achieving its objectives. Chapter 7 of *The Green Book* states that an evaluation is "…conducted in the same manner as an economic appraisal…" with almost identical

[9] *A New Deal for Trunk Roads in England.* (1998). Department of the Environment, Transport and the Regions.

procedures. The air quality MCDA reported here is concerned only with appraisal, so that will be the focus of our comparisons, below.

3.3 Comparisons of MCDA with CBA

At first glance, the two approaches seem quite different: eight steps for MCDA, only four for CBA. On closer examination, many of the sub-headings in CBA's step 3 appear as separate steps in MCDA. In particular, Step 3 in CBA is similar to Steps 3 through 6 in MCDA. Taking a more holistic view, each emphasises somewhat different aspects of appraisal. This can be seen more clearly in Table 3. We have tried to find language that would be acceptable to both practitioners of MCDA and CBA. Note that the similarities cover obvious aspects of any appraisal, however it is done. The order of those seven steps is also common to both approaches. We have included 'Identify objectives' and 'Identify options' on the same row because both approaches would agree that either might be done before the other.

Table 3: Similarities and differences in the reported steps for MCDA and CBA.

<table>
<tr><th colspan="2">Similarities between MCDA and CBA</th></tr>
<tr><td colspan="2">Establish context and justification for action.
Establish aims and rationale for the analysis.
Consider the context and scope of the appraisal. Is government intervention warranted?</td></tr>
<tr><td>Identify objectives and criteria
Organise criteria (outputs against which options will be appraised) and objectives in a hierarchy.</td><td>Identify options
List possible policies, strategies or actions to achieve the objectives.</td></tr>
<tr><td colspan="2">Appraise the options
Value the costs and benefits of the options on the criteria.</td></tr>
<tr><td colspan="2">Derive an overall valuation
MCDA: Calculate overall weighted scores. CBA: Calculate the net benefits or costs.</td></tr>
<tr><td colspan="2">Examine results
Identify the 'best' option.</td></tr>
<tr><td colspan="2">Conduct sensitivity analysis
Consider the impact of alternative scenarios and changes in key variables.</td></tr>
<tr><th colspan="2">Features unique to MCDA</th></tr>
<tr><td colspan="2">Design the socio-technical system for conducting the MCDA.
Describe the consequences of the options.
Check the consistency of the scores on each criterion.
Assign weights for each of the criteria to reflect their relative importance to the decision.
Calculate overall weighted scores at each level in the hierarchy.
Look at the advantages and disadvantages of selected options, and compare pairs of options.
Repeat the steps until a 'requisite' model is obtained</td></tr>
<tr><th colspan="2">Features unique to CBA</th></tr>
<tr><td colspan="2">Discount costs and benefits to 'present values', when appropriate.
Adjust for differences in tax between options.
Adjust for risk and optimism.
Determine the affordability of options.
Devise implementation plans.
Present the results.
Implement the solution.
Track the success of the policy, programme or project in achieving its objectives.</td></tr>
</table>

At this level of description, MCDA and CBA appear very similar. They share seven main steps for carrying out an appraisal.
Examination of the features unique to each shows considerable differences, but practitioners of the two approaches will immediately recognise steps they include implicitly. For example, in commercial work, a financial value objective is usually included in an MCDA, and realised in a net present value criterion. This would require the discounting of all costs and benefits to present values to reflect the decision maker's time preferences. Similarly, from a CBA perspective, describing the consequences of each option is surely done early in the project, as is a degree of consistency checking of all valuations.

Another example is in the handling of uncertainty. Although not mentioned in the MCDA steps of Table 1, The MCA Manual includes in section 6.3 a discussion of how uncertainty might be handled, with an appendix devoted to one of the ways. Both CBA and MCDA acknowledge the usefulness of decision trees. Traditionally, a decision tree has been the model of choice for an uncertain future, and MCDA for modelling multiple objectives in situations of certainty. In practice, both uncertainty and multiple objectives characterise most appraisals. Since MCDA is an extension of decision theory, incorporating uncertainty is not a problem. It is done through the use of probabilities, which permits application of the expected utility rule, scenarios built into the value tree, risk criteria or sensitivity analysis on the scores or weights. All these approaches are mentioned in *The Green Book*.

Further consideration of Table 3 shows a higher prescription of detail for CBA than MCDA, not surprisingly given that an economic valuation is the focus. MCDA is less prescriptive in how appraisals are carried out for each objective. It might well include a financial valuation based on a CBA, along with other criteria. However, without a structure such as that of MCDA, the bringing together of all the criteria is necessarily informal. Non-monetisable criteria are simply 'considered' (see Step 3.8 in Table 2) alongside those which are monetisable. MCDA provides a formal analytical framework for this stage in the decision making process. Research in behavioural decision theory shows that 'informal' trade-offs typically fall short of the 'best' results obtained by a formal trade-off model[10].

Table 3 hints at differences in process, which became more apparent to us during the course of the project. According to *The Green Book*, 'the first step in appraisal is usually to carry out research, to identify the scope of the issues involved and the basis for government action'. The research gives the information that provides a justification for action. MCDA begins by designing the appraisal process itself, which may or may not require research.

A serious difference between MCDA's use of the term 'objectives' and the way they are treated in *The Green Book* can be the source of considerable confusion between the two approaches. Section 4.1 of *The Green Book* states: "If an intervention seems worthwhile, then the objectives of the proposed new policy, programme or project need to be stated clearly." Note that 'objectives' refer to the desired consequences of government intervention. The next sentence says: "This allows the identification of the full range of alternative options which government may adopt." Thus, options may be defined in terms of government objectives, a potential confusion of options with objectives. In MCDA, objectives that are the consequences of governmental intervention would normally be included, but so, too, would other objectives, such as those associated with the means for achieving the objectives, or the differing objectives of stakeholder groups, or unintended objectives. *The Green Book* says very little in Chapter 4, Setting Objectives, about how objectives are to be identified. On the other hand, *The MCA Manual* offers practical advice on the topic in section 6.2.5, and further practical detail can be found in Keeney's extensive treatment of value-focussed thinking[11]. That said, additional guidance beyond *The Green Book* is provided by various publications from the Cabinet Office's Regulatory Impact Unit.

[10] Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge: Cambridge University Press.
[11] For example, Keeney, R. L. (1992). *Value-Focused Thinking: A Path to Creative Decisionmaking*. Cambridge, MA: Harvard University Press.

Our enquiries suggested that a typical appraisal process is to form a project team which includes CBA practitioners and specialists in the topic under consideration. Participants meet periodically, gathering data and analysing it between meetings, perhaps commissioning further research, until a solution is developed that identifies the best option. Consultation with external specialists and those affected by the proposals may take place at any time, but it is required after the best option has been identified so that this solution can be further refined. So, *The Green Book* acknowledges the importance of the social process, but this only emerges at Step 4.3, Table 2, which is in the implementation stage.

The MCA Manual, on the other hand, is explicit about the importance of consultation throughout all the eight steps. This is evident in Step 1.2, Table 1, 'Design the socio-technical system for conducting the MCDA', in which the social and technical aspects of the appraisal are designed together at the start of the project. The Manual suggests facilitated workshops[12] as one approach: groups of key players at any stage in the appraisal meet to resolve the issues at that point in the appraisal, guided by an impartial facilitator who is a specialist in MCDA and group processes, and supported by modelling on-the-spot using computer programs designed for multi-criteria analysis. The London School of Economics version of this is called *decision conferencing*, which has been developed over the past 20 years and is widely used by decision analysis practitioners and organisations in both public and private sectors. Decision conferencing provides a structure for consultation at any or all stages of an appraisal, so when the appraisal is complete, implementation usually follows as a matter of course since all key players were involved during the analysis. This provides transparency to the model, provides an audit trail of all assumptions and judgements, makes it easy to present the results up the organisational hierarchy, and achieves the buy-in of those involved. Decision conferencing was used on this project for the three workshops exploring the application of MCDA to air quality policy.

Another difference in process is that *The MCA Manual* discusses the vexed issue of how much modelling is enough: just sufficient to resolve the issues at hand. It recommends a 'rough-and-ready' MCDA at the start of any project, largely to enable sensitivity analyses to be carried out to identify those areas in which more modelling effort should be devoted. *The Green Book* is largely silent on this crucial topic, usually only recommending more analysis if it can increase confidence in the final results.

It is probably fair to say that CBA involves a great deal of 'back room' modelling, so that subsequently presenting the results may require careful explanation of how those results were obtained. Published case studies using MCDA typically involve working with groups of key players, drawn from practitioners, experts, stakeholders, and interest group representatives, in short, anyone with a perspective on the issues at hand who could contribute to resolving them. Some versions of MCDA involve 'back room' modelling, but the decision conferencing process is mainly 'front room' modelling, carried out during the meeting, the model projected so it is fully visible to all participants.

However, all the above discussion about process tends to obscure the most obvious difference between MCDA and CBA. All cost-benefit analyses express valuations of costs and benefits in monetary units. Multi-criteria analyses use value-scaling techniques to ensure that all valuations are made with non-monetary value scales whose units are equated so the scales can be combined. CBA expresses all valuations in monetary units, using assumptions about markets where these exist, or engaging in willingness-to-pay (WTP), willingness-to-buy (WTB), or willingness-to-accept (WTA) studies, or any other non-market valuation technique when markets do not exist or are imperfect, as is the case for many heath and environmental benefits that arise from reductions in air pollution. Whether or not all or none of the criteria are monetisable, by whatever methods, MCDA can proceed, but it introduces two formal concepts that are foreign to CBA: value functions and criterion weighting. Value functions translate input data, such as performance measures, into non-monetary values. They allow for the possibility of non-linearity in the value of a particular output measure (e.g., reducing an undesirable health effect by half may not double the value of that reduction). This non-linearity is implicit in CBA. Criterion weightings express the trade-offs between criteria, indicating how much an increase in value on one criterion corresponds to an increase in value on another. Value functions and criterion weights are illustrated later in this report, so a short description here will suffice to introduce the concepts.

[12] See section 6.2.2 of *Multi-Criteria Analysis: A Manual.*

Consider two criteria that Keeney[13] found to be of concern to Los Angeles residents about air pollution: public health and safety, and cost of an air pollution abatement program. Where possible, criteria in MCDA are defined as objectively as possible. Here, public health and safety was defined as the annual number of otherwise healthy adults diagnosed as having a 20% impairment in lung function, and cost was defined as the total annual cost. Good and bad levels of each criterion were established: no adults diagnosed as having impaired lung function was considered good, while 1,000 diagnoses was deemed bad; zero cost was good, $7.2 billion annual cost was bad. Criterion weights were assessed for individuals affiliated with nine (of 11) stakeholder groups. The key question asked of participants was "How much additional annual cost would be acceptable to change the level of the criterion from bad to good?" One participant, for example, said it would be worth $100 million to eliminate 1,000 cases of impaired lung function. Later, as a check, the same person was asked how much cost would be acceptable to eliminate one adult case of impaired lung function. To be consistent, the person should say $100,000. These redundant questions are typically employed by MCDA practitioners to provide different ways of thinking about the same issue and as a way to identify and reconcile possible inconsistencies. This participant gave a unit of value for health and safety as $100,000, while further questioning established that a unit of program cost was $1,000,000. Thus, the value units are weighted in the ratio of 1:10. The final weight is then taken to apply to the range from bad to good; it is called a 'swing weight', for it expresses the increase in value in moving from the least to most preferred positions on the scale, here, bad to good. The direct assessment of swing weights was employed in the three workshops for the current project.

Inconsistencies in the judgement of unit values can arise because value functions are not linear, in which case a swing weight for the whole range is obtained. Although Keeney did not explore his participants' value functions, it is possible that some people would consider the amount to be spent to reduce one adult's impaired lung function depends on the total number of affected adults. If few people are currently affected, perhaps a higher cost per person would be judged acceptable than if a great many people are now affected. Of course, some might argue in the opposite direction: lower costs per person the more that are affected. Non-linearity of the value associated with the number of cases of impaired function is handled in MCDA with a value function constructed by asking people to judge the incremental increases in value associated with stepped reductions from Bad to Good in the number of people affected.

This discussion raises the issue of whose values are to be included in an MCDA. That depends on who is commissioning the MCDA and for what purpose. The Los Angeles air quality study was sponsored by the Electric Power Research Institute "to develop a methodology for involving stakeholders in decisions that affect them." Meetings were held with between one and three people from eleven different groups concerned about air quality to determine their concerns and develop concrete objectives. After each meeting, the objectives were structured in a value tree and sent back to the participants for revisions. Eventually, all objectives from all groups were combined into one model showing health and safety concerns, quality-of-life concerns and economic and equity concerns. It is interesting to note that air pollution levels themselves were not included in residents' concerns, so it could have been fruitful to examine the consequences of actions to control air pollution on the criteria that are the main concerns of citizens; a role for back-room modelling. It would then have been useful to bring the participants together to explore further why the objectives were of concern, to score the policy options on the criteria, to weight the criteria and to explore the sensitivity of the overall results to differences of opinions about disputed scores and weights.

An example of how MCDA can help to resolve differences between competing stakeholder groups is a study by Gardiner reported in von Winterfeldt and Edwards[14]. The Coastal Zone Development Act of California requires an appraisal of every development proposal and a subsequent approval or disapproval. Commissioners are often in conflict with developers, so the MCDA made explicit the individual value functions and weightings for eight criteria. Gardiner used the model to explore differences between conservationists and developers, and found considerable agreement for many proposals. It appears that when opposing stakeholders meet, they quickly focus on their differences, negating the effect of the many

[13] Ibid., pp. 317-325.
[14] Ibid., pp. 279-282.

criteria on which there is agreement. The MCDA provides a more balanced way to ensure that all criteria enter the appraisal, with the result that overall differences are not as great as they seem in an unstructured, face-to-face meeting. Opposing sides then find it easier to make concessions.

The important point about value functions and criterion weightings is that they are necessarily subjective concepts, requiring human judgement for their determination. Indeed, appraisals in MCDA can be made directly on value scales in the absence of performance measures. However they are arrived at, they are made explicit and open, and so are subject to public scrutiny. The many judgements made in a cost-benefit analysis are typically less obvious, less public and more technical. The route from an objective performance measure to a value to a weighted value to a final result is clear in MCDA. The route in CBA from a performance measure to a monetary unit can be opaque, as we learned during the course of this project.

3.4 Summary

We found many similarities between MCDA and CBA in *what* should be done when conducting an appraisal, but found substantial differences in *how* to do them. Both approaches agree about the seven steps that constitute an appraisal, with *The Green Book* additionally concerned about implementing the solution. MCDA practitioners typically see their role as analysts, leaving implementation to the client.

Differences in how an appraisal is to be carried out include MCDA's ability to trade off any collections of objectives against one another, including both monetary and non-monetary objectives, whereas in CBA the non-monetisable benefits are 'considered' at the end of the analysis. Many differences between the approaches boil down to differences in process, not substance or content. A CBA begins with research; an MCDA often starts with a 'rough-and-ready' MCDA model. Both CBA and MCDA analyses are usually carried out by a project team, with more 'back-room' modelling in CBA than in MCDA. Consultation of experts, specialists, stakeholders and others is routine throughout an MCDA, often with groups of key players in facilitated workshops. This helps to create ownership of the results by the client, who then implements the appropriate solution. In *The Green Book*, implementation is seen as part of the appraisal process, which ends with presenting the results in a draft report that is widely circulated for consultation. This technically oriented approach to implementation is in contrast to MCDA's view of it as more of a social process requiring engagement of key players, including the client, during the modelling.

The most striking difference between MCDA and CBA is in the way valuations of costs and benefits are expressed: monetary units for CBA, non-monetary units for MCDA. To use non-monetary units in MCDA requires judgements of value functions, which translate objective performance measures into values, and assessments of criterion weights, which express the trade-offs between values on the various criteria. Judgement is, of course, required for both approaches, which raises the question of how the approaches bring a degree of objectivity to their analyses. This is a deep issue, for which the following is only an inadequate summary.

In CBA, valuations are based on market prices, but when these are not available, they are based on research in which stakeholders' valuations are obtained through revealed preference or stated preference techniques. Revealed preferences are those inferred from participants' choice behaviour, while stated preferences are obtained by questionnaires and interviews that elicit the stakeholders' willingness to pay (WTP) for desired outcomes or willingness to accept (WTA) particular outcomes. These approaches bring a degree of external validation to CBA because the research is conducted independently of those contributing to the decision at hand.

Similar approaches are found in MCDA, as illustrated by the LA air quality and the California Coastal Commission examples of the previous section. In these cases, stakeholders were consulted and their values elicited, using a variety of techniques, many of which are identical or similar to WTP and WTA studies. Like CBA, MCDA results in a value model, which contains both qualitative and quantitative elements. As Keeney[15] states, "The general procedure for building a value model is essentially the same as

[15] Ibid., Section 5.1, p 130.

for any model. … A value model should be developed from first principles, sound logic, reasoned judgments, and carefully acquired, consistent data." In this sense, we see no difference between CBA and MCDA as regards their objectivity. However, MCDA models are often constructed using directly-assessed judgements. To ensure the realism and consistency of these judgements, several steps are taken. First, assessments are often obtained from groups of stakeholders or specialists, who are chosen to represent all the key perspectives on the issues. Working in groups allows peer review to take place during the elicitation process, often resulting in judgements that are better than could be obtained from even the best expert. Second, where possible, data relevant to those judgements are collected and provided to participants. Third, proper scaling techniques, based on sound, empirically tested principles, are used, such as psychophysical scaling techniques[16]. Fourth, the impartiality of the facilitator of these groups can often identify the unhelpful effects of 'group think' and he or she can then challenge the group to justify their views backed with data. Fifth, extensive consistency checking helps to identify inconsistent judgements, which can then be resolved through group discussion. Finally, 'gut checks' throughout the process and sensitivity analyses at the end, help to resolve any remaining sense of unease about the results, sometimes by further modifications to the value model, and at other times through new insights gained from the analysis.

Keeney also identifies as a key issue "whose values should be utilized in the problem." His answer is that "the value models of any stakeholder interested in a particular decision context are appropriate. In some decision contexts, it is fruitful to construct multiple value models for the different stakeholders. After all, by definition, each of the stakeholders cares about the consequences, so finding out which consequences they care about and why may be useful[17]."

Keeney goes further in arguing for public involvement in any problem that affects the public. "Simply put, the impacted public should be asked what set of objectives they want to be used for creating alternatives and for evaluating the alternatives. To do this, stakeholders and/or specifically created groups of the public to represent "the public" may be useful. Once a good set of objectives is determined, then governmental staffs or various interest groups can proceed with additional analysis[18]."

In summary, we see no substantive difference between the potential for developing valid CBA or MCDA models. The revealed preference and stated preference techniques of CBA are used in MCDA, but the latter more frequently uses stated preference elicitation methods in group settings.

[16] See von Winterfeldt and Edwards, Chapter 7, Value and utility measurement.
[17] Ibid. Section 5.6, p. 152.
[18] Ibid. Section 5.6, pp. 152-3.

4. Study Approach

The project followed the steps given in Table 1 for conducting an MCDA. We began by meeting the project sponsors, DEFRA and EA, and other key civil servants interested in MCDA, on 18 October 2002 to learn more about the project's objectives and scope, to discuss the EU and UK government policy context, to learn about current UK work on air quality and to discuss a work plan. Three objectives were agreed:

1. to gain a better understanding of MCDA
2. to understand the differences between MCDA and CBA
3. to show what additional benefits MCDA would bring to the appraisal process

An overview of MCDA presented at the start of the meeting, and the questions that arose from the presentation, began the process of achieving mutual understanding. Participants agreed that at this stage only a pilot study was required to achieve the objectives. We proposed a case study approach for illustrating the application of MCDA methodology to appraising UK air quality policy options because a 'live' example would take us through all the stages of an MCDA, whereas a retrospective study would only deal with the technical aspects of applying an MCDA.

Subsequent work by the sponsors identified possible case studies related to the Air Quality Strategy objectives (specifically the work which reviewed the particles objectives in 2001) and protection of ecosystems (specifically the current AQS objectives with respect to ecosystems and protection offered to SSSIs). A kick-off meeting for the MCDA study on 9 January 2003 began the process of seeing how MCDA could be applied to air quality policies. Interviews in January with key sponsors of the project provided insight into the objectives for both the air quality and ecosystem MCDA models, and began the process of developing specific benefit and cost criteria from the objectives. The two case studies were completed in three workshops conducted in a manner similar to a decision conference: participants representing a diversity of perspectives on the issues were guided by impartial facilitators through the construction of a computer-based model that was projected throughout the process so all participants could see what was going on at any stage. The decision conferencing process is designed with three objectives in mind: to generate shared understanding of the issues, to create a sense of common purpose, and to gain commitment to action. All these purposes can be achieved without obtaining consensus on every issue.

The steps shown in Table 1 were followed with slight differences between the workshops to accommodate the differing requirements of the models developed. Information relevant to the topic was sent to participants prior to the first meeting. After introductions, subject matter experts provided briefings on the topic under consideration. The lead facilitator gave a brief introduction to decision conferencing and MCDA modelling, and the group then engaged in agreeing or developing the options to be considered. For the air quality study, the options were grouped into three areas, one set for London, one for Scotland, and one for the rest of England and Wales.

The group discussed policy objectives and agreed both cost and benefit criteria, drawing on their experience and on published work. The facilitators often questioned participants to see why suggested criteria were considered important, and to help the group develop criteria that were comprehensive, clearly defined, related to objective performance measures where possible, did not double count, and were mutually preference independent[19]. For the ecosystems case study, the criteria were structured as a hierarchical value tree.

Participants were then engaged in scoring the options on the criteria. Scoring always required participants to make comparisons between pairs of options on a given criterion. Typically, the most preferred option on a criterion, as judged by the group, was given a preference value score of 100, and the least preferred a score of zero. This resulted in a relative, or interval, scale, on which differences in value between pairs of

[19] Two criteria are considered to be mutually preference independent if the valuations of options on one criterion can be made without knowing the valuations of the options on the other criterion.

options can be compared. As a result the ratio of scores for two options cannot be meaningfully interpreted in terms of value because the zero point on the scale does not represent 'no value'. Ratios of differences can be interpreted meaningfully, but not ratios of the temperatures themselves.

For some criteria, objective performance measures were entered into the model, but these were always converted linearly into 0-100 value scales, directly when larger performance measures were considered preferable (more lives saved is preferable to less), and inversely when lower performance measures were judged to be more preferred (less pollution is preferred to more). No non-linear value functions were required for either case study. In one case, for the ecosystem study, a qualitative scoring technique, known as 'Macbeth scoring' was used. This required only verbal judgements of the magnitude of the value difference between pairs of options for a give criterion, constantly checked by the Macbeth software tool for consistency. If inconsistencies were detected, the software flagged the problem and suggested ways to change the judgments to achieve consistency. It then converted the verbal judgements into a set of value scores that preserved all the input paired comparisons. The two case studies provided opportunities to employ different scaling techniques.

Weighting the criteria followed. Since all preference values were expressed on relative 0 to 100 scales, the units of value on the scales were not necessarily equal. (Both Celsius and Fahrenheit scales show 0° to 100° portions, but a Celsius degree represents a larger increment in temperature than a Fahrenheit degree.) Weighting equated the units of value. The process, swing weighting, asked participants to judge the swing in value from the least to most preferred position on one scale as compared to the swing on another. First, the options associated with 0 and with 100 on one scale were identified. Next, the options on another scale at the 0 and 100 positions were identified. The group was then asked to judge how big the swing from 0 to 100 is on one scale as compared to the swing on the other scale. The operative question was, "How big is the difference between the least and most preferred option on this scale, and how much do you care about that difference, as compared to the difference on the other scale, and how much you care about it?" After a weight of 100 was assigned to the criterion whose swing was judged to be the largest, participants then compared each criterion's swing to the one weighted 100, resulting in weights for all the criteria. Since differences in values were being compared, consistency checks were made of the ratios of the weights. Also sums of weights were required to be consistent. For example, for three criteria weighted 100, 60 and 40, the swing from 0 to 100 on the first criterion must equal the sums of the swings from 0 to 100 on the other two criteria (swinging from 0 to 100 on the second criterion gives 60% as much value as the swing on the first criterion, and the swing on the third criterion provides 40% as much value, so the two together, 60 plus 40, must give the same value as the 100 swing on the first criterion).

With all scores and weights input to the associated computer program, the results were then calculated and displayed to participants. Discussion of the results was followed by sensitivity analyses which explored the effects on the overall results of imprecision in the inputs and differences of opinion. Time permitting, each workshop finished with participants reflecting on their experience of the meeting.

All workshops were facilitated by Larry Phillips and Adrian Stock of Catalyze. Although air quality strategy occupied the first two workshops, and ecosystems the third, we present the ecosystems work first because that model is a more transparent introduction to MCDA.

The remainder of this report uses the two case studies, along with our experience of using MCDA in the public sector, to add to the MCDA/CBA comparisons of the previous section, and to achieve the objectives of the project. Shaded boxes appearing throughout the remainder of this report summarise our findings related to the project's objectives. Rather than leave these findings to a separate section, we felt it more appropriate to flag the issue immediately in the context that gave rise to the finding.

Reports detailing the inputs and results of the workshops are available separately from this report.

5. Ecosystems model

The idea behind this workshop was to appraise SO_2 and NOx policy options that could offer increased ecosystem protection to SSSIs (Sites of Special Scientific Interest) on AQS objectives for ecosystems and vegetation. Since time at the workshop was limited, it was agreed to focus only on the SO_2 objectives. In part, this case study was chosen because it concerned ecosystem impacts which are difficult to value with CBA. The six participants at the workshop, held on 19 March, were drawn from DEFRA, JNCC, EA and SEPA to represent a diversity of perspectives. Information relevant to the protection of SSSIs was circulated in advance of the meeting. An initial briefing by the DEFRA (AEQ) participant explained the direct effects on plants of noxious emissions, the policy framework, current strategies for NO_x and SO_2, the number of SSSIs affected and criteria currently considered for appraising policy options.

5.1 Options

After a brief introduction by the lead facilitator to MCDA modelling, the group agreed to look at policy options for sulphur dioxide (SO_2) in its effects on SSSIs. An Entec study[20] had identified five possible policy options, all relating to 2010:

1. 20μg/m^3 annual mean concentration with 100% of SSSIs in compliance, measured as a mean across the whole site.
2. 10μg/m^3 annual mean concentration with 100% of SSSIs in compliance, measured as a mean across the whole site.
3. 10μg/m^3 annual mean concentration with 99.5% of SSSIs in compliance, measured as a mean across the whole site. This option represents the expected outcome of policy measures currently in place and is referred to as the status quo (SQ).
4. 20μg/m^3 annual mean concentration with 100% of SSSIs in compliance, measured as a maximum for any point on the site.
5. 10μg/m^3 annual mean concentration with 100% of SSSIs in compliance, measured as a mean across the whole site.

5.2 Objectives and criteria

The group extended the initial discussion of criteria, making it possible to construct a hierarchical 'value tree', Figure 1, showing policy objectives at higher levels and appraisal criteria at the lowest level.

Finding 1: Language confusion. The language used by practitioners of MCDA and CBA does not always coincide. The Entec report presents the above policy options as 'objectives' in the sense that they are the objectives the policy is intended to achieve, both the threshold concentration and the per cent compliance. Table 3 of the Entec report shows both policy options and consequences of the policies, which initially we found confusing as to which was which. The report also refers to 'available abatement measures', which we took to mean physical measurements of the consequences of the policy options, but actually means those actions that can be taken to reduce SO_2 concentrations. Thus, options were confused with objectives, and courses of action with physical measures. Other confusions continually arose during the course of the project (e.g., the use of 'scenarios').

One contribution of decision analysis is its consistent use of language: options, events, consequences, objectives and criteria take very specific meanings. In MCDA an objective is a general statement of a desired end state at some time in the future, and a goal is a measurable objective. Here, the objective is to reduce SO_2 at SSSIs, and the above five policy options are expressed as goals, or targets. This is an important distinction because options are under the control of the decision maker—the government can eventually choose one of the policy options—whereas the consequences of the chosen option are not entirely under the government's control. Thus, it is important to distinguish between inputs and outputs.

Whereas MCDA has a well defined taxonomy, this does not apply to CBA, nor indeed to economics in general. This provides more freedom for CBA, and could pose a problem in MCDA when a particular situation does not fit with the decision analysis taxonomy. However, experience suggests that the taxonomy of decision analysis is far more often helpful than inappropriate.

[20] *Protection of Sites of Special Scientific Interest under the Air Quality Strategy for England, Scotland, Wales and Northern Ireland*, DEFRA, Final Report, Entec UK Limited, in preparation.

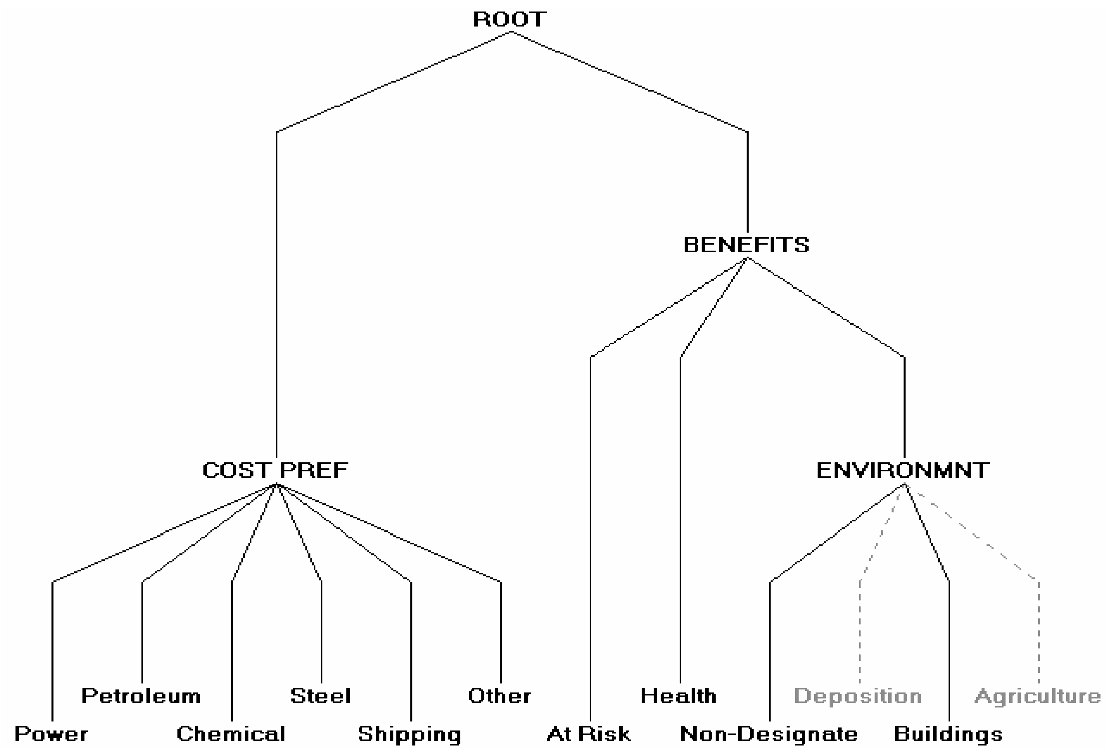


Figure 1: The ecosystems value tree. The Root node represents the overall appraisal of the options, taking account of all the bottom-level criteria.

Cost criteria, measured in annual £millions, identified which industry would likely have to pay for new environmental measures. The following sectors were identified:

- **Power**
- **Petroleum**
- **Chemical**
- **Steel**
- **Shipping**
- **Other**

COST PREF simply means Cost Preference, a reminder that later displays will show each option's preference for overall costs (lower costs give higher preference scores), not the costs themselves.

Benefit criteria were split between SSSI risk, human health and environmental benefits.

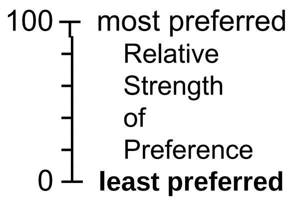
- **SSSI At Risk**: Extent to which damage to ecosystems and vegetation (specifically those plants that provided the reason for the SSSI designation) will be reduced at a SSSI.
- **Health Effects of Air Pollution**: Extent to which reductions in SO_2 concentration levels are expected to reduce the health effects of air pollution, specifically in terms of a reduction in deaths brought forward and respiratory hospital admissions per year.
- **Non-Designated SSSI Benefits**: This criterion represents the benefit to flora and fauna at a site that is secondary to the reason the site was designated an SSSI.
- **Deposition**: This measures the reduction in deposition of Sulphur. This has been ignored for the pilot study as it has a very small effect.
- **Building Damages**: The reduction in damage to buildings due to the reduction of SO_2, which is a key pollutant associated with materials erosion and hence impact on buildings.
- **Agriculture**: There were no data available for agriculture effects so these benefits have not been quantified. This criterion was included because SO_2 can influence crop yield or quality.

Considerable discussion attended the definition of the SSSI At Risk criterion and how it related to the number of sites that complied with the policy. Some group members believed that plants were at risk by definition if the site did not meet regulations. However, the definition above was finally agreed to show that even compliant sites might suffer damage risks, especially if the policy was measured as a mean across the whole site. The SSSI At Risk criterion refers only to those plants which led the SSSI designation, so this criterion is preference independent of the Non-Designated criterion.

Finding 2: Completeness of the objective set. The Entec study is a cost-effectiveness study, so naturally it focuses just on the costs of meeting the objectives set down in the policy options. It also looks at Health and Buildings effects, which are analysed separately. By contrast, the MCDA is capable of capturing all the relevant criteria simultaneously. At Risk and Non-Designated criteria were not included in the Entec study; both these criteria represent additional benefits to flora and fauna.

5.3 Scoring the options

The purpose of scoring in MCDA is to achieve for each criterion a scale extending from 0 to 100, with those points on the scale clearly defined. Those points can be defined independently of the options under consideration, or with reference to the least and most preferred options, as was done for the ecosystems study. Scoring is the process of assigning numbers to the options on the criterion under consideration. The scale can be thought of as representing relative value or 'strength of preference'. Since it is an interval scale, like Celsius or Fahrenheit, whose zero points and units of measurement are arbitrarily chosen, it is important to appreciate that only differences in value can be compared on such scales. For example, if options A, B and C have been scored at 100, 80 and 0, respectively, then the difference in preference value between B and C is four times as big as the difference in value between A and B. It is wrong to suggest that option B, scoring 80, is 80% as good as option A, scoring 100, or that option D, scoring 40, is half as good as option B. It is ratios of differences in the scores that can be compared, not ratios of the scores themselves.



In this study, two approaches were used, direct and indirect scoring. The Entec study provided cost information for all cost criteria, but for only two options on the Health criterion. Direct scoring required the group to identify the most and least preferred options on a given criterion, assign these scores of 100 and 0, then score the remaining options so that differences in the scores reflected differences in preference. This approach was taken for two of the benefit criteria, Non-Designate and Buildings.

Finding 3: CBA can provide inputs to MCDA. All cost information was provided by the Entec study. In principle, CBA could also provide valuations of benefit criteria that are capable of monetisation. It then becomes possible in the MCDA to see the effects of trade-offs among all the criteria, which can give a different result from looking first at only monetary criteria, then adjusting these judgmentally by 'taking account' of non-monetary criteria.

Indirect scoring was used for all the cost criteria. The input cost data were rescaled (by Hiview, the software used to model the MCDA): for a given cost criterion, the least costly option was assigned 100, the most costly a 0, and all the others between in proportion to their costs. Indirect scoring was also used for the Health benefit criterion. Input data, number of deaths and hospital admissions avoided per year, were available for two options; figures for the other two options were judged by the group. The group then judged that these input data could be converted by an inverse linear function to preference values.

The group found it difficult to give direct numerical scores to the At Risk criterion, so an indirect, qualitative approach was taken: Macbeth scoring, using a software tool for qualitative MCDA. Macbeth created a matrix whose rows and columns were the options. Participants' judgements described the extent to which each row option was more preferred than the column option, using verbal descriptors of extreme, very strong, strong, moderate, weak, very weak and no preference. If agreement could not be reached about a particular comparison, more than one verbal descriptor was entered. From this matrix of verbal descriptors, Macbeth was able to create a set of scores for the options that were entirely consistent with the qualitative judgements. The resulting 0-100 value scale was transferred into Hiview.

Finding 4: MCDA can be both a social and technical process. The initial version of the MCDA model included a criterion defined as the number of compliant sites. As participants started scoring the policy options on this criterion it became evident that the criterion rewards the less stringent policy options: the less stringent the option, the larger the number of compliant sites. But these higher scores for less stringent policies were contrary to participants' preferences. Discussion moved to recognising it is the number of sites that remain at risk that is the important criterion: less stringent policy options would leave a greater number of sites at risk, resulting in the policy receiving a lower preference score on the criterion. After much discussion, the SSSI At Risk criterion was agreed; it captures the extent to which damage at a site will be reduced as a consequence of the policy option. This redirection of thinking was not just the result of the MCDA, it also gained much from the presence of several participants with different viewpoints, demonstrating the socio-technical nature of decision conferencing.

In summary, a mixture of scoring techniques was used for the cost and benefit criteria.

Criterion	**Scaling technique**
All cost criteria	Inverse linear (the higher the costs, the lower the preference value).
SSSI At Risk	Macbeth qualitative scoring.
Health	Inverse linear after group judgements of three options were given compared to data provided for the other two options.
Non-designate; Buildings	Directly assessed preference values.

Finally, it is worth noting that many other techniques could have been used, including any of the approaches used in CBA. Data, empirical studies, modelling, agreed standards and human judgement are all employed in MCDA. In this study, the collective judgements of specialists played a prominent role. The reliance on expert judgement in this workshop is based on studies of experts, which show that experts can provide outputs that are as good as, and in some cases superior to, explicit models[21]. The justification for working in groups is based on experience with decision conferencing and recent research showing that the consensus judgement of a group can be better than even its best member[22].

Finding 5: MCDA provides an analytical structure for comparing monetary and non-monetary outputs. Box 10 in *The Green Book* shows the valuation techniques available for CBA: real or estimated market prices, or willingness to pay studies, using revealed preference, hedonic pricing or stated preference techniques. In Annex 2, 'Valuing Non-market Impacts', *The Green Book* provides helpful guidance for criteria that are not easily monetisable. However, it also admits that many criteria will still be very difficult for CBA. Chapter 5, paragraph 5.76 states, "Costs and benefits that have not been valued should also be appraised; they should not be ignored simply because they cannot easily be valued. All costs and benefits must therefore be clearly described in an appraisal, and should be quantified where this is possible and meaningful." Extending the valuation techniques and using the framework of MCDA can realise a broader, meaningful appraisal.

Note the distinction between 'valued', which means valued in monetary terms, and 'quantified', which means 'non-monetary' measures. In MCDA, 'valued' takes on a broader meaning; 'value' is capable of being expressed in non-monetary units. *The Green Book* notes that "the most common techniques used to compare both unvalued costs and benefits is weighting and scoring (sometimes called multi-criteria analysis)", and refers to the *MCA Manual*. However, from the perspective of MCDA, 'weighting and scoring' are just two steps in an eight-step process. The treatment in *The Green Book* is incomplete, and no suggestion is made that MCDA and CBA could be integrated.

---

[21] Shanteau, J. (1999). Decision making by experts: The GNAHM effect. In J. Shanteau & B. A. Mellors & D. A. Schum (Eds.), *Decision Science and Technology: Reflections on the Contributions of Ward Edwards* (pp. 105-130). Boston/Dordrecht/London: Kluwer Academic Publishers.

[22] Regan-Cirincione, P. (1994). Improving the accuracy of group judgment: A process intervention combining group facilitation, social judgment analysis, and information technology. *Organizational Behavior and Human Decision Processes, 58*, 246-270.

5.4 Weighting the criteria

Weighting in MCDA is the process of ensuring the equality of a unit of preference value on all the 0 to 100 scales. The scoring process results in a relative scale for each criterion, but the value difference between 0 and 100 may be different for each scale, as a Celsius degree represents a different unit of temperature from a Fahrenheit degree. The process of equating the units of value was accomplished by asking participants to compare the swings in preference from 0 to 100 on all the scales. This is a process of identifying the options associated with 0 and with 100 on a particular scale, then asking the group how big the difference is between those options and how much they care about that difference, as compared to 0-100 differences on other criteria. The process is called 'swing-weighting'[23].

It is this step that is perhaps the most misunderstood in MCDA, for weights are often thought to reflect the absolute importance of the criteria. Not so. The following example was given in the workshop. If you were to purchase a new car, would you consider price to be important to your decision? Most people answered, 'yes'. Since you can't consider all possible cars on the market, you construct a short list that includes just five cars. Suppose the difference in cost between the least and most expensive is £200. Now is price an important consideration in your decision? Most participants said 'no'. But if the difference in price were £2,000, then many said that would be more important. Unless, we pointed out, you are very wealthy, in which case a difference of £2,000 might not have much impact. The point of this example is two-fold: first, a criterion's weight depends on the range of difference in the input data, and secondly, on how much you care about that difference. Inevitably, that has to be a judgement.

In the Los Angeles air quality example on page 14, the swing weight was first judged between bad and good extremes. For the ecosystems workshop, the extremes were determined by the most and least preferred options on a given criterion. The advantage of this approach is that hypothetical extremes, such as bad and good, don't have to be defined. Instead, the extremes are established by well-defined options. However, the disadvantage is that if a new option is brought into play, and it extends the existing range on a criterion, then that criterion's weight should be increased. Fortunately, MCDA software anticipates these problems and provides several techniques to ensure the consistency of weights when ranges change.

Weights in MCDA represent trade-offs. Once the weights are established, they show how much an increase on one criterion is equal to an increase on another. An increase of 9° Fahrenheit is equal to an increase of 5° Celsius. Some methods for assessing weights rely on making judgements of trade-offs, or by comparing the best on criterion A and the worst on B with the worst on A and the best on B, but this latter method doesn't work when those hypothetical combinations are physically impossible. The more general technique, swing-weighting, is generally the preferred approach; it was used here. First, within a cluster of criteria, the largest swing is identified. It is given a weight of 100. Then all other swings on criteria within the cluster are compared to 100 and assigned appropriate weights. Thus, if the swing on another criterion is judged to provide half the swing in preference value, then that criterion is given a weight of 50. Weights are ratio scale numbers, since they compare differences, so there need not be a zero-weighted criterion.

The weights for the cost criteria in the ecosystems model were judged by the group to be proportional to the differences in the costs. Essentially, the group agreed that an expenditure of a pound for one payer would be equal to a pound for another. We appreciate that this may seem an obvious assumption for those who see money as fungible. But there are exceptions: an amount of consumption may not be equated to the same amount of public expenditure, and in many commercial applications, a pound spent from one budget does not necessarily equal in non-monetary value a pound spent from another budget. However, for this study a pound is a pound, whoever spends it. The rationale for weighting the cost criteria under this assumption is illustrated in Figure 2.

[23] A brief introduction to swing weighting is given in Section 6.2.10, pp. 62-3, of the MCA Manual.

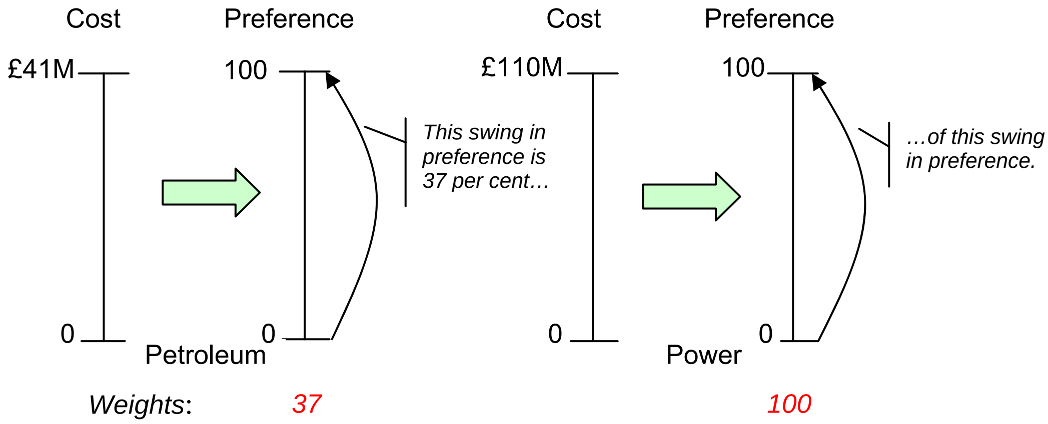


Figure 2: Weighting the cost criteria. The most costly option for both the Petroleum and Power industries was option 5, 10µg/m³ annual mean concentration with 100% of SSSIs in compliance. The least costly was the Status Quo, option 3, which implies no additional cost. For the petroleum industry, the cost was assessed at £41 million per year, and for the power industry, £110 million per year. When those input scales are transformed to 0-100 preference scales, the swing in value for the petroleum industry is 37% as large as the swing for the power industry. Thus, the weights for the preference scales are in the ratio of 37 to 100.

Benefit weights are more complex to assign because their units of benefit are different. First to be weighted were the benefits under the Environment node. The group were asked, "Of all the swings within these criteria, which represents the biggest difference that you care about?" The swing for Non-Designated benefits was judged the largest swing in value and was assigned a weight of 100. The benefit swing for Buildings was judged to be half the value so was assigned a weight of 50. The Deposition and Agriculture criteria were weighted at 0 as they had no scoring data and would not affect the model.

The next weighting question was asked higher up the tree at the Benefits node. Because the environmental benefits had already been weighted, we needed only to consider the criterion that was weighted 100 under that node, Non-Designated. The group was asked, "Of the swings in At Risk and Health, which represents the biggest difference that you care about?" The largest swing in value was agreed to be SSSI At Risk and this was assigned a weight of 100. Health swing was judged only one-tenth as important so was weighted at 10. Next, the swing on At Risk was compared with the highest-rated swing of the two criteria under the Environment node, Non-Designate. At Risk was judged to be the larger swing, with Non-Designate assessed at 55. Back at the Environment node, the weight assigned to Non-Designated was adjusted to 55, which meant that the Buildings' criterion weight had to be reduced by 55%, resulting in a weight of 27.5. Weights at this level were now summed and entered as 82.5 at the Benefits node. This systematic process resulted in the following relative benefit weights (the bottom row shows the weights normalised so they sum to 100; in the final model all these weights would then be halved to reflect equal weight on the Cost and Benefits nodes):

At Risk	Health	Non-Designate	Buildings
100	10	55	28
52	5	28	15

With costs and benefits weighted separately, the next step was to equate a unit of cost with a unit of benefit. The group discussed viewpoints of various stakeholders including conservationists, DEFRA, EA, Treasury and Health. It was agreed to value units of cost and benefit equally for simplicity. This completed the model.

Finding 6: Who weights the criteria? In MCDA, human judgement is required to establish the swing in value between the 0 and 100 reference positions on a given criterion as compared to the swings on other criteria. The resulting weights are scale constants that equate the value of a unit change on one criterion to a unit change on another. Here, the group did not find it particularly difficult to assess weights for the benefit criteria, though each weight was the subject of discussion and negotiation within the group. But they did find assessing weights for Costs and Benefits difficult, so settled on a simple compromise of 50-50.

Our experience shows that if the group is composed of knowledgeable experts, assessing weights, a task that is unfamiliar and at first difficult, becomes manageable and elicits further insights into the problem at hand. However, when we phrased the question as "How big is the difference between least and most preferred options, and how much do you care about it?", the group asked who should be doing the caring. Since the topic is government policy, this raises the question of who can represent the public interest. Presumably, the answer is the government minister responsible for the policy. Others argued that it is the affected public who should care, though nobody argued that the payers' caring should be solicited. Clearly, this is a serious question of how the public interest should be represented in weighting the criteria in an MCDA. Further discussion and debate about this issue is certainly required.

Confronted with this problem in the past, we have recommended that after a thorough sensitivity analysis to see which weights matter to the overall result, several sets of weights be constructed to represent different interests. A Treasury perspective could be simulated with more weight on Costs than Benefits, a Green perspective represented by more weight on Benefits than Costs and more weight on the At Risk and Non-Designate criteria, and a human health perspective simulated by placing more weight on Benefits and Health. If the same option ends up overall most preferred under all three perspectives, then a firm recommendation can be made to a minister. If not, then the minister can be advised of the sensitivity of the recommendation to different perspectives, and the sensitivity analyses can guide the minister's civil servants in drawing up conditional recommendations, leaving the minister to make the final judgement call.

5.5 Results

The computer software calculated weighted averages of the preference values, starting at the bottom of the tree and moving to the top. The final result is shown in Figure 3.

The Root Node

ROOT Data Breakdown

BRANCH	Wt	20mn100	10mn100	SQ10mn99	20mx100	10mx100	CumWt
COST PREF	50	99	77	100	59	0	50.0
BENEFITS	50	15	76	12	31	100	50.0
TOTAL	100	57	76	56	45	50	100.0

Figure 3: Overall results for the ecosystem model, assuming equal weights on costs and benefits. Overall, 10mn100 is the most preferred option with a weighted value of 76.

The five options form the column headings, with overall preferences for costs in the first row (higher numbers reflecting lower costs), and for benefits in the second row. Since those two objectives are equally weighted, the Total row shows the weighted average of costs and benefits. It is clear that options 1 and 3, 20mn100 and SQ are very similar, but that option 2, 10mn100, is overall preferred.

The bar graph representation in Figure 4 shows this very clearly. Option 2 contributes more benefits than either 1 or 3, but the greatest benefit is from option 5, 10mx100. Unfortunately, this option is also the most costly.

Further insight can be gained by viewing the overall cost and benefit preference values for each option as a data plot, Figure 5, which is, of course, not affected by the difficult judgement that a unit of cost is equal to a unit of benefit. There the advantage of option 2 becomes apparent. Although it isn't as high in benefits as option 5, nor as low in cost as the SQ, it is relatively high in benefits and low in costs. The upper surface of the green portion of the plot forms an 'efficient frontier', created by four options. Only option 4 is inside the frontier, showing that it is dominated by option 2, which is less costly and more beneficial. Thus, option 4 could be eliminated from being put forward as an efficient policy option (remembering that this is a trial model, not a final statement).

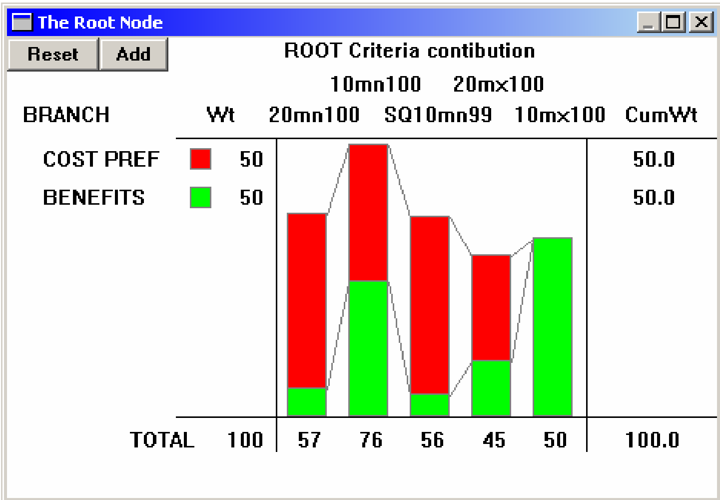


Figure 4: Overall results for the ecosystem model showing a graphical breakdown of the overall values into their cost and preference values. Longer red bars indicate lower costs, while longer green bars show greater benefits.

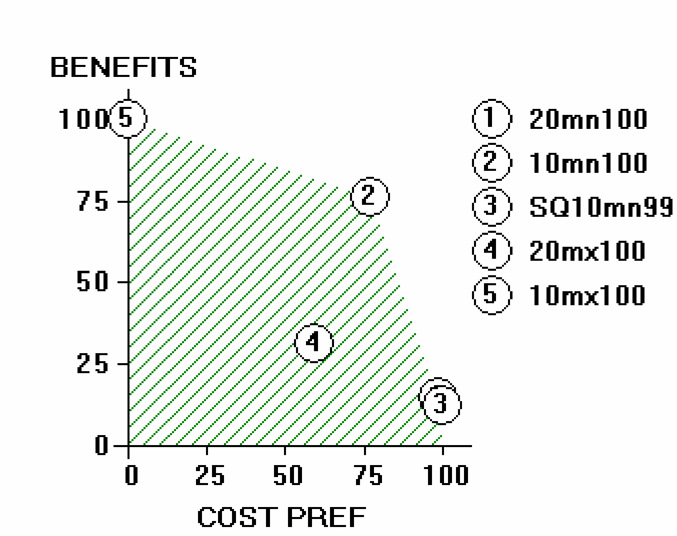


Figure 5: Plot of the overall benefit preference values versus the overall preference values for costs.

Finding 7: MCDA graphs aid understanding. There are just two figures in the Entec report, neither associated with showing overall results of the cost-effectiveness analysis. In fact, the reader of that report will search in vain for a recommendation. Table 5, the summary information of the cost-effectiveness solutions for the policy option for SO_2, leaves the reader still wondering which option is best.

By contrast, the above results, the first showing that option 2 is overall most preferred, taking both costs and benefits into account, the second giving a clearer picture of why option 2 is most preferred, provided a clear bottom line for the group. Seeing that option 4 is dominated was another insight that wasn't obvious in the table of numbers in Figure 3.

5.6 Sensitivity analyses

The position of option 2 toward the upper right portion of Figure 5 indicates that it will remain the overall most preferred option over a wide range of relative weights on costs and benefits. This can be seen clearly in Figure 6.

The vertical red line at 50 shows the current relative weight on costs. That weight is varied over its entire range from 0 to 1.0, with the resulting overall scores for each option given by the slanting lines (note that options 1 and 3 are virtually identical). The shaded portions indicate weights for which the best option would change from option 2. Thus, if the weight on Cost were reduced to about 25, indicating no concern for cost, then option 5 would be most preferred. Increasing the weight to about 75 would lead to maintaining the status quo. So, whether a unit of benefit is considered three times more value than a unit of cost, or vice versa, option 2 would remain the most preferred. The overall preference for option 2 is very robust to imprecision or differences of opinion about the relative weight on costs and benefits.

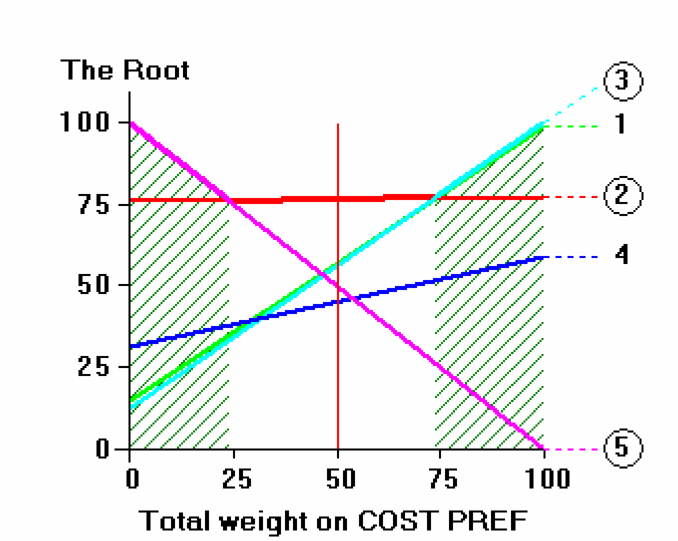


Figure 6: Sensitivity analysis on preference for costs showing the robustness of option 2 to wide changes in that weight.

Health benefits were of particular concern to some in the group, so a sensitivity analysis was conducted on this criterion's weight. It is shown in Figure 7. Clearly the preference for option 2 is unaffected by that weight over a very large range for the weight.

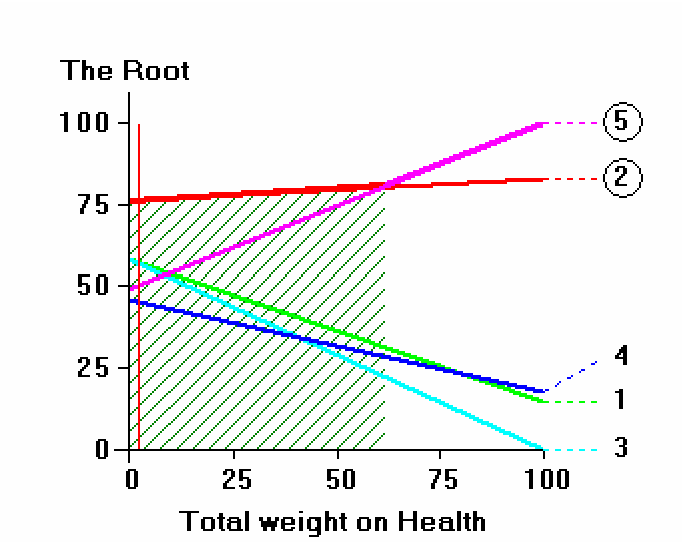


Figure 7: Sensitivity analysis on the Health criterion's weight. Note the cumulative weight of about 3% on that criterion has to be increased to capture about 62% of the total weight in the model before option 5 becomes most preferred.

Finding 8: Precision in weights is not required in MCDA. But we don't know ahead of time which weights are critical and which aren't. So precise values are given to all weights (though they can be assessed qualitatively using Macbeth weighting), and then are followed by sensitivity analysis to detect the key ones. Sometimes the overall result shows great sensitivity to particular weights. However, this is not common. Far more often, substantial changes in the input data, value functions (linear is usually a 'good-enough' approximation, as here for the Health criterion) and weights make little difference to the position of the most preferred option, as can be seen in the two sensitivity analyses. This reasons for this insensitivity, which is largely a consequence of the structure of MCDA models, is explored in some depth by von Winterfeldt and Edwards[24].

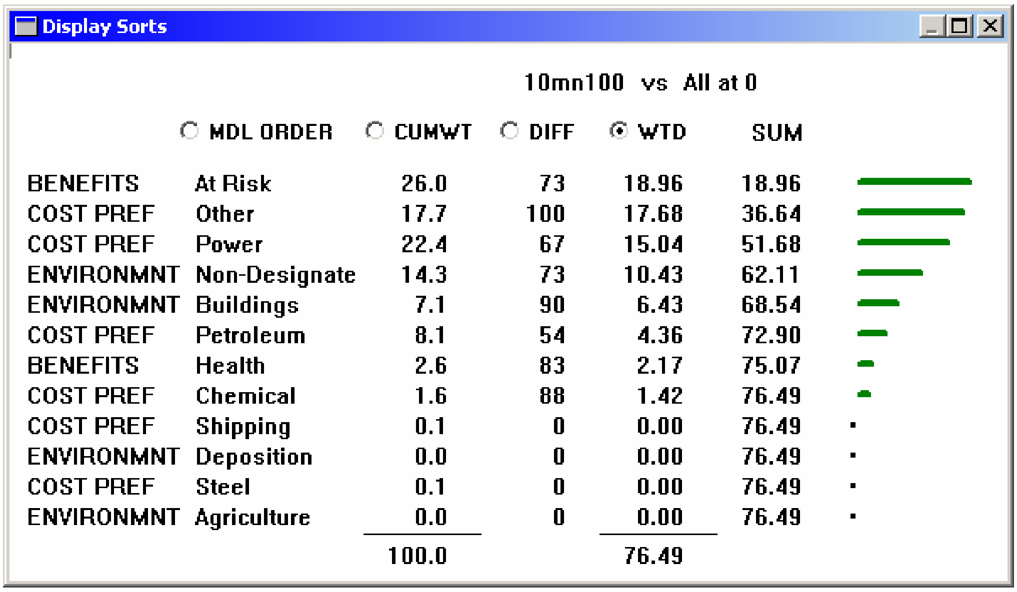
A good example of this principle can be seen in the ecosystems model. With all input data and weights in place, it is possible to monetise all the non-monetary criteria using as a reference the weights and range of scores on a monetary criterion. The weights represent trade-offs, so other units can now be turned into monetary units. A convenient criterion for this treatment is the Health criterion, which represents the numbers of deaths or admissions to hospital prevented by the policy option. A few simple calculations, using the cumulative weights of the Power and Health criteria, reveals that at the current 50-50 trade-off between costs and benefits, one death/admission prevented is worth £30,000. With a 25-75 trade-off, the figure is £90,000, and with weights of 75-25, £10,000. So over the range of weight on costs from 25 to 75, all of which show option 2 to be overall most preferred, the value of reducing one death/admission ranges from £90,000 to £10,000. Our experience is that some people find assessing weights makes more sense to them with swing weighting questions, whereas others like to see trade-offs in monetary units.

5.7 Advantages, disadvantages and comparisons of the options

Further insight can be obtained from an MCDA model by conducting 'sorts', which order the options in any one of three ways: to show advantages of a particular option, to show its disadvantages and to compare it with another option.

An option's advantages are indicated by high overall preference values on criteria with greater weights. An example for the most preferred option 2 is shown in Figure 8.

[24] See chapter 11, "Sensitivity analysis and flat maxima" in Winterfeldt, D. vW., & Edwards, W. (1986). *Decision Analysis and Behavioral Research.* Cambridge: Cambridge University Press.



Display Sorts

10mn100 vs All at 0

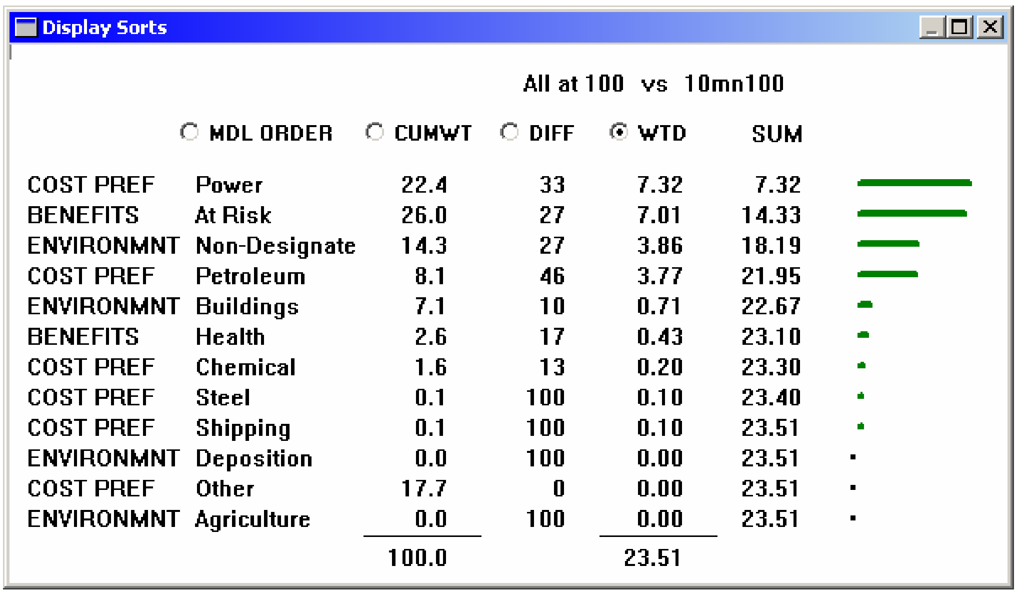
		CUMWT	DIFF	WTD	SUM
BENEFITS	At Risk	26.0	73	18.96	18.96
COST PREF	Other	17.7	100	17.68	36.64
COST PREF	Power	22.4	67	15.04	51.68
ENVIRONMNT	Non-Designate	14.3	73	10.43	62.11
ENVIRONMNT	Buildings	7.1	90	6.43	68.54
COST PREF	Petroleum	8.1	54	4.36	72.90
BENEFITS	Health	2.6	83	2.17	75.07
COST PREF	Chemical	1.6	88	1.42	76.49
COST PREF	Shipping	0.1	0	0.00	76.49
ENVIRONMNT	Deposition	0.0	0	0.00	76.49
COST PREF	Steel	0.1	0	0.00	76.49
ENVIRONMNT	Agriculture	0.0	0	0.00	76.49
		100.0		76.49	

MDL ORDER ○ CUMWT ○ DIFF ◉ WTD

Figure 8: Advantages of option 2, 10mn100.

The first numerical column, CUMWT, gives the cumulative weight (the product of the normalised weights from the top of the value tree to the relevant criterion) on each of the 12 criteria. The next column, headed DIFF is the difference between the preference value of that option on each criterion from zero; in other words, it is that option's preference value. Note that the option scores 100 on 'Other', whose cumulative weight is fairly high, and it scores well, 90, on 'Buildings' whose weight is lower. The next column, WTD, shows the weighted preference value, represented by the green bars. Now the weighted preference values on Other and Buildings are quite different. The SUM of all weighted preference values is the cumulative sum, 76.49, which is the overall score of option 2 shown in Figure 4 rounded to 76. The numbers in the WTD column are interpreted as the 'part scores' which together make up the overall score. Thus, only the first three criteria contribute about two-thirds of the total value of the option. The main advantages are the reductions in deaths and hospital admissions and its relatively low Other and Power costs.

Now look at option 2's disadvantages, Figure 9: low scores on heavily weighted criteria.



Display Sorts

All at 100 vs 10mn100

MDL ORDER ○ CUMWT ○ DIFF ◉ WTD

		CUMWT	DIFF	WTD	SUM
COST PREF	Power	22.4	33	7.32	7.32
BENEFITS	At Risk	26.0	27	7.01	14.33
ENVIRONMNT	Non-Designate	14.3	27	3.86	18.19
COST PREF	Petroleum	8.1	46	3.77	21.95
ENVIRONMNT	Buildings	7.1	10	0.71	22.67
BENEFITS	Health	2.6	17	0.43	23.10
COST PREF	Chemical	1.6	13	0.20	23.30
COST PREF	Steel	0.1	100	0.10	23.40
COST PREF	Shipping	0.1	100	0.10	23.51
ENVIRONMNT	Deposition	0.0	100	0.00	23.51
COST PREF	Other	17.7	0	0.00	23.51
ENVIRONMNT	Agriculture	0.0	100	0.00	23.51
		100.0		23.51	

Figure 9: Disadvantages of option 2, 10mn100.

Here the DIFF column shows the option's score subtracted from a hypothetical option that scores 100 on all criteria. Here, just four criteria, shown by the long green bars, are identified as the main disadvantages. Note that two, possibly three, of them are also advantages. That occurs when an option's score is neither very high nor very low. If the criterion is fairly heavily weighted, then the same criterion can be both advantage and disadvantage. This time the cumulative sum is the difference of the overall score from 100: 100 – 76.49 = 23.51.

Finally in what ways is option 2 better or worse than the status quo? Figure 10 gives this comparison. The DIFF column gives the difference between each option's scores on the relevant criterion, so the WTD column is the weighted difference. Now two clear advantages of option 2 over the status quo emerge: its reduction in damage to ecosystems and vegetation at SSSIs, and the benefit to flora and fauna at a secondary site. It is interesting to note that these are really the only two comparative benefits; the health effect is very small. The red bars show just two advantages of the status quo over option 2, its lower costs (zero) to the petroleum and power industries.

Display Sorts

10mn100 vs SQ10mn99

		MDL ORDER / CUMWT	DIFF	WTD	SUM
BENEFITS	At Risk	26.0	73	18.96	18.96
ENVIRONMNT	Non-Designate	14.3	73	10.43	29.39
BENEFITS	Health	2.6	83	2.17	31.56
ENVIRONMNT	Buildings	7.1	5	0.36	31.91
ENVIRONMNT	Deposition	0.0	0	0.00	31.91
COST PREF	Other	17.7	0	0.00	31.91
ENVIRONMNT	Agriculture	0.0	0	0.00	31.91
COST PREF	Shipping	0.1	-100	-0.10	31.81
COST PREF	Steel	0.1	-100	-0.10	31.71
COST PREF	Chemical	1.6	-13	-0.20	31.51
COST PREF	Petroleum	8.1	-46	-3.77	27.74
COST PREF	Power	22.4	-33	-7.32	20.42
		100.0		20.42	

Figure 10: Comparison of the 10mn100 overall best option with the status quo.

One of the main uses of sorts is to help a group think more deeply about possible new options. When a group is challenged to think of ways to obtain more benefits on just the key benefit criteria and lose cost on the key cost criteria, but worse on the cost and benefit criteria with low weights, they often become inspired to find 'win-win' options, which move closer to the upper right hand corner of Figure 5.

This completed the final steps of exploring the model.

Finding 9: MCDA provides methods for discovering the key advantages and disadvantages of an option, and the important ways it differs from other options. These 'sorts' can stimulate a group to think creatively and generate new options that are better in the benefits that matter, and lower in the key costs, resulting in 'win-win' options. In CBA comparisons are also made between different options, but because monetary units are used throughout, the comparisons are more transparent than for MCDA where both scores and weights have to be taken into account.

5.8 Reflections
Participants were invited to reflect on this workshop. In no particular order, these comments were received (here paraphrased):

- This got me to think more deeply about the issues.
- I was more involved.
- Can get lost in the detail of a CBA.
- I like the sensitivity analyses.
- It would be useful to incorporate some risk scales.
- Could choose an appropriate dose/response curve and incorporate it in a value function.
- Sensitivity analyses were important and useful.
- Good to line up the important factors. The sorts are a useful aid to thinking.
- We need to see how this can be used and briefed upward.
- In any study, this would be helpful to use to structure the study and prepare us mentally.
- This was a more useful demonstration than the air quality workshops.
- Good to talk early about all the different impacts.
- Sensitivity analysis is great.
- I'm still not entirely clear about the At Risk criterion.

5.9 Discussion
This model demonstrates the use of MCDA for appraising a set of options against any collection of monetary and non-monetary criteria. Many of the steps are similar to those encountered in CBA. Three major differences that did not appear in comparing *The Green Book* with the *MCA Manual* are now apparent: first, the MCDA was conducted in a group setting, second, human judgement was explicitly incorporated into the model, and third, extensive use of projected, graphical displays lent clarity to understanding model results.

Although MCDA can be conducted solely as a back-room modelling process, most MCDA practitioners prefer to work with groups of key players representing diverse perspectives on the issues. Working in a facilitated group enables participants to exchange ideas and engage in rational argument, which can lead to changes in attitudes and to new insights. There is clear research evidence that the consensus judgements of a group can be better than the judgements of even the most knowledgeable person in the group, or of the average of individual judgements[25]. The conditions for this superior performance include impartial facilitation of the group's work, application of a model to structure thinking about the issues and use of information technology to ensure that the model-building process is clearly visible at all times. This process of collectively exploring a difficult issue and formulating a new product was very evident in the discussion of the SSSI At Risk criterion.

Finding 10: Conducting an MCDA with groups of key players can lead to products that would not have been obtained with back-room modelling. Insofar as a CBA is conducted by a project team who consult specialists and experts, either individually or in small homogeneous groups of participants, with a draft report circulated for consultation and finally revised in an attempt to accommodate the differing responses, the overall result will not be the same as an MCDA. It cannot be, for inadequate opportunity for conflicting perspectives to thrash out their differences is offered with back-room modelling and consultation. Furthermore, implementation will suffer because only the project team owns the final report. In a decision conferencing application of MCDA, all key players are engaged from the start and they all meet periodically in workshops to move the work on to the next step. The model is built in their presence, with full transparency. By the end of the process, participants feel a sense of ownership of the results and can easily brief them upward. Implementation usually proceeds quickly.

This finding raises an important question about how civil servants work. Our sponsors, who arranged the workshops, found that participants were often unwilling to commit an entire day of their time to one activity, let alone the two days that are typical of decision conferences. We see no way around this; it

[25] Ibid., Regan-Cirincione, P.

simply isn't cost-efficient to arrange half-day workshops—the group just gets warmed up to its task when it is time to stop. But we have experience of working in the public sector with a series of one-day workshops, interspersed with back-room data gathering, which works quite well. It becomes less effective if new people keep turning up, requiring time to be spent briefing them. A series of workshops, held on non-consecutive days, attended by the same participants each time, can work very well.

Incorporating human judgement in an MCDA model raises questions of objectivity and validity. Many judgements have to be made in the process of conducting any appraisal, CBA or MCDA. Appraisals are always about future impacts, and no data are available about the future, so human judgement is an inescapable element in any appraisal. MCDA makes explicit rather more of those judgements than CBA.

That raises the next issue of the validity of human judgements of strength of preference. This can be a contentious issue, some people arguing that valuations should not be based on strength-of-preference methods; they should instead be derived from actual or hypothetical choices, which are objective and observable. We agree with von Winterfeldt and Edwards[26]that such data "can be used to infer underlying structures that may or may not predict choice behavior. How well either kind does so is an empirical question; the evidence we know suggest that both do well." The research base for their view is extensive, carried out by behavioural psychologists[27] and economists. This research recognises the importance of proper assessment techniques, in particular methods developed many decades ago in the discipline of psychophysics (the relationship between an individual's sensations and the objectively-measurable stimuli that produced them). In this study, sound difference scaling techniques were used, followed by consistency checks.

Finding 11: Proper scaling techniques are required in MCDA. The benefit scales created for the ecosystems model were interval scales, which are characterised by arbitrarily chosen zero points and units. But ratio scales, with a fixed zero point (zero means zero amount of the property being measured), are also used in MCDA, as they are in CBA. Understanding the difference between these scale types is essential in eliciting the scale values and interpreting results correctly.

The third issue, extensive use of projected, graphical displays, is part of a deeper issue—the environment in which the group workshop is held. One of us has written elsewhere[28] about the impact of the environment on the group; we consider this to be an oft-neglected aspect of an MCDA. The point is quite simple: the environment in which a group works can affect the quality of the products. Two simple principles guide us in setting up a room for a workshop: every participant should be in direct eye-to-eye contact with all others, and material displayed on whiteboards, screens and flipcharts should be visible and readable by everyone. We prefer square rooms and an arrangement of chairs and tables such that no straight rows of participants are created. This arrangement ∩ works better than this one: ⊓. A data projector located at the open end of the arrangement provides the means for showing the computer's display, and whiteboards or flipcharts on either side of the screen enable the facilitator to capture key points. A continuous supply of drinks and light refreshments should also be available in the room

Unfortunately, all of the workshops were conducted under less than these ideal environmental conditions. Small rooms, an excess of furniture and immovable rectangular tables prohibited us from creating a more satisfactory environment.

Finding 12: MCDA modelling in workshops requires a supportive physical environment. We prefer off-site meetings in good surroundings, for then participants give their full attention throughout the day. Provision for the costs of these meetings should be included in the project's budget.

---

[26] Ibid., p. 211.

[27] See, for example, Fischhoff, B. (1991). Value elicitation: Is there anything in there? *American Psychologist, 46*(8), 835-847. Reprinted in D. Kahneman and A. Tversky (Eds.), *Choices, Values and Frames*, Cambridge: Cambridge University Press, 2000. Fischhoff provides a table of 'Conditions Favorable to Articulate Values', many of which are met in decision conferences. Also see Chapter 7, "Value and utility measurement" in von Winterfeldt and Edwards.

[28] Phillips, L. D. (1989). People-centered group decision support. In G. Doukidis & F. Land & G. Miller (Eds.), *Knowledge-based Management Support Systems.* Chichester: Ellis Horwood.

5.10 Summary
The ecosystems MCDA model examined five policy options which varied in the severity of target performance for SO_2 concentration, measured as a mean or maximum, and compliance in their effect on SSSIs. Available information provided costs for the options, broken down into six categories of payers. Benefits were assessed by the group under four criteria, SSSI At Risk (extent that damage to plants is reduced), Health Effects of Air Pollution (number of deaths and hospital admissions avoided), Non-Designated SSSI Benefits (secondary benefits to flora and fauna) and Building Damages (reduction in damage to buildings). Weighting the criteria enabled a computer to calculate overall weighted preference values for the options, with the result that option 2, 10μgm/m3 annual mean concentration with 100% of SSSIs in compliance, measured as a mean across the whole site, was shown to be overall most preferred. Extensive sensitivity analyses showed that considerable vagueness or disagreement about weights would lead to this same conclusion. (Of course, no policy recommendations should be made on the basis of this incomplete model.)

Twelve findings relevant to the comparison of MCDA with CBA emerged from this part of the study:

1. Different uses of language between MCDA and CBA can cause considerable confusion.
2. MCDA can capture any set of criteria, monetary and non-monetary.
3. CBA can provide the monetary inputs to MCDA; they are not alternative approaches.
4. MCDA combines social and technical processes.
5. MCDA provides an analytical structure for comparing monetary and non-monetary outputs.
6. In MCDA, human judgement is required to establish relative weights of the criteria.
7. MCDA graphs, which are typical outputs, aid understanding.
8. Precision in weights is not required in MCDA.
9. MCDA provides methods for discovering the key advantages and disadvantages of an option, and the important ways it differs from other options.
10. MCDA's focus on group modelling can lead to different results from CBA conducted in the 'back room'.
11. Proper scaling techniques are required in MCDA.
12. MCDA modelling workshops require a supportive physical environment.

Overall, The final model demonstrated how MCDA handles a mix of financial and non-financial objectives, creating an overall preference ordering of the options, establishing an efficient frontier consisting of the best options for a given cost, and enabling extensive sensitivity analyses to be carried out to see the effects on the overall ordering of imprecision in the input data and differences of opinion about judgements of value. In addition, it showed how to establish the key advantages and disadvantages of selected options, and how to find the important ways in which pairs of options differ.

6. Air quality models

The two air particles workshops were conducted over two days, 3 February, and 12 March, both attended by 10 to 12 people, plus the authors acting as facilitators.

A 'top-down' model of policy options in different parts of England and Scotland was only partially completed during the first day. Further work on health benefits was the subject of the morning at a second workshop on 12 March. In the afternoon, the group created a 'bottom-up' model of a portfolio of potential policy measures for lowering PM_{10} concentrations. With many new participants at the second workshop, much time was devoted to discussion of other aspects of the model as well as health benefits, and that model was not completed. Sufficient work was done on the PM_{10} model to show how MCDA modelling of either policy options or measures can be combined to realise an overall portfolio of options that maximises value-for-money.

The structure of these models is very different from the hierarchical value tree created for the ecosystems workshop. Whereas that model is most appropriate for appraising or evaluating several options on many criteria, the air quality models show how to design the best collection of options from several areas or budget categories, creating best portfolios of options. This type of model can be used for prioritisation and resource allocation. If an R&D organisation is seeking to create a portfolio of options across many products in development, this MCDA model will prioritise all possible options on a value-for-money basis. If each of many organisational units creates several proposals for next year's budget, the model can find the best combination of options across all the units, with the result that some units may find their budgets increased while others are decreased. In general, this MCDA approach seeks to design an overall system from options associated with several sub-systems. It is sufficiently general that it has been used by many UK Local Authorities to determine how best to cope with budget cuts, to help an organisation's board find the best portfolio of strategies for its various operating divisions, to allocate a limited budget for advertising several different products, to prioritise research proposals for support by a government research unit, and even to find an affordable design configuration for a new warship. It is useful to think of this approach as a way of putting together separate appraisal models into one overall model that generates an efficient frontier. The intention is to help decision makers take decisions that are more equitable across different budget categories, hence the name of the supporting software, Equity, used in these workshops.

This approach is sufficiently different from the value-tree model that it warrants revision of Table 1, shown on the next page. The best way to understand the steps is to illustrate them with the air quality case studies. Both models will be explained in the steps to follow.

6.1 Context

The Air Quality Strategy for England, Scotland, Wales and Northern Ireland, published in January 2000, established standards and objectives to by met between 2003 and 2008 for eight key air pollutants. Two subsequent documents[29], "…explain the changes that the Government and the devolved administrations propose to make to the Strategy's objectives to take account of the latest health evidence," and they provide the economic analysis to back up proposed changes. While not proposing any new national measures, these reports conclude that if the objectives for air particles were strengthened, "…the implied cost per added life year which they would produce is in no way excessive."

In attempting to see how the NAO recommendation to use MCDA as an aid to appraising possible policy options for air particles could be implemented, the groups at the two decision conferences recognised that a single policy might not be appropriate for the whole of the UK. A target would be most difficult to meet in London, while relatively easy in Scotland. Thus, a package, or portfolio, of policy options should be examined, recognising that not all regions could with equal cost-effectiveness meet a single standard.

---

[29] DEFRA. (2001). *The Air Quality Strategy for England, Scotland, Wales and Northern Ireland.* London: Department for Environment, Food & Rural Affairs. DEFRA. (2001). *An Economic Analysis to Inform the Review of the Air Quality Strategy Objectives for Particles.* London: Department for Environment, Food and Rural Affairs. These will be referred to as 'supporting studies' in subsequent sections.

Table 4. A summary of the MCDA process for prioritisation and resource allocation

Applying MCDA for prioritisation and resource allocation

1. **Establish the decision context.**
   1.1. Establish aims of the MCDA, and identify decision makers and other key players.
   1.2. Design the socio-technical system for conducting the MCDA.
   1.3. Consider the context of the appraisal.

2. **Identify the areas and the options to be appraised for each area.**
   2.1. Identify budget categories, or areas.
   2.2. Generate options within each area (if appropriate, include current investments extending into next year's budget, as well as new proposals).

3. **Identify objectives and criteria.**
   3.1. Identify criteria for assessing the consequences of each option. These usually cover costs, benefits and risks.
   3.2. Define the criteria by including sub-criteria that capture all the forward costs, benefits and risks.

4. **"Scoring". Assess relative preference scores for each option.**
   4.1. Describe the consequences of the options.
   4.2. Assess costs for all options.
   4.3. Assign relative scores (values) that express the extent to which the options in an area achieve the objectives represented by a given criterion.
   4.4. Check the consistency of the scores on each criterion—compare differences of scores for interval scales, and ratios and sums of scores (balance-beam checks) for ratio scales.

5. **"Weighting". Assign swing weights for each of the criteria to reflect their relative importance to the decision.**
   5.1. Assign within criterion weights that equate the units of value from one area to the next on a given criterion
   5.2. Assign across criterion weights that equate units of value from one criterion to the next.

6. **Combine the weights and scores for each option to derive an overall value.**
   6.1. Calculate total forward costs for each option.
   6.2. Calculate the total risk-adjusted benefits for each option.
   6.3. Divide the risk-adjusted benefit for each option by its total cost to give the value-for-money ratio that establishes the priority of the option relative to other options.

7. **Examine the results.**
   7.1. If necessary, reorder the options in each area (by increasing cost for mutually exclusive options, and usually by the value-for-money ratio for cumulative options).
   7.2. Examine each area's graph of cumulative risk-adjusted benefit versus cumulative cost.
   7.3. Examine the overall efficient frontier.
   7.4. Examine the current portfolio's position relative to the efficient frontier.
   7.5. Establish a budget figure to define a frontier portfolio. Note which projects fall outside the frontier portfolio.

8. Conduct sensitivity analyses.

8.1. Explore the effect of different scores for individual options on their position in the efficient frontier, in particular for selected options that fall outside the frontier portfolio.

8.2. Explore the effects of an area's within criterion weights by changing the weights.

8.3. Try different weighting systems for the across criteria weights to identify those projects that consistently remain high or low priority whatever the across criteria weights.

8.4. Keeping an eye on the frontier portfolio, attempt to improve the current portfolio to move it as close as possible to the frontier portfolio.

8.5. Examine the effects of trading into the frontier mandated or required options that fell outside the frontier.

8.6. Repeat the above steps until a "requisite" model is obtained.

An additional concern was raised by the participant from the Environment Agency, which is charged with ensuring that Local Authorities are taking reasonable steps to meet the air quality standards. For the EA, Government policy is given; their concern is with enforcing measures that can be taken at local level to meet the standards. So their concern is to achieve a better understanding of the most cost-effective collection of measures. This is a 'bottom-up' approach of finding the best portfolio of abatement measures.

Over the course of the two workshops, both concerns were addressed. We have elected to discuss both in parallel in the following sections, hoping that the comparison between 'top-down' and 'bottom-up' approaches will enhance understanding of how MCDA can contribute to the construction of best portfolios of options.

6.2 Areas and Options

Policies model

For the first workshop, we used the supporting studies to prepare a draft model to present to the group. It consisted of three areas, the geographical areas that the core team suggested should form the focus of the model:

1. UK excluding London and Scotland
2. London
3. Scotland

Within each of these areas, several policy options were identified (called 'levels' in Equity), some being fairly lax, others quite stringent as regards air quality standards expressed as a 24-hour mean in micro-grams per cubic meter, an annual mean and number of days the standard is exceeded. Thus, a policy abbreviated as 50(35/yr), 32, represents a 24 hour mean of 50μgm/m^3, not to be exceeded more than 35 days/year with an annual mean of 32μgm/m^3.

Discussion by participants revealed that most of these policy objectives were provided by European Union Directives on air quality. EU directives are expressed in terms of both annual mean and occurrences of exceedence to allow for the different ways that member states control air quality. This setting of two standards caused some confusion for the group: should the focus be on one of these as a proxy for both, or both at the same time?

After some discussion, the group decided that policy options should be expressed in terms of the annual mean as this has more bearing on the health benefits of reduction in PM_{10} concentrations. Two consequences of this decision followed:

1. An extra option was added to hit the EU target of not greater than 35 days/year exceedence of the 50µg/m^3 24 hour mean. This correlated to a 32µg/m^3 annual mean.
2. A new benefit criterion was created: number of days the 50µg/m^3 24 hour mean was exceeded.

The resulting model structure is shown in Figure 11.

Level	UK ex LON & SCOT	LONDON	SCOTLAND
5	50(7/yr), 20an	50(7/yr), 20an	
4	20 annual mean	50(10-14/yr),23-25an	50(7/yr), 18an
3	23-25 annual mean	20 annual mean	18 annual mean
2	50(35/yr), 32an	23-25 annual mean	23-25 annual mean
1	SQ: 50(80/yr), 40an	SQ: 50(35/yr), 40an	SQ: 50(35/yr), 40an

Figure 11: The model structure for the air quality policy model.

The options at level 1 represent the current policy. Options above level 1 represent other possible policy options the group considered feasible. The options are, of course, mutually exclusive for each of the three geographic areas: one and only one can be chosen for each geographical area, though it isn't necessary for the same policy to be chosen across the areas.

Measures model
We prepared a draft model based on the six major sources of PM_{10} given in the supporting studies:

1. Domestic Combustion
2. Quarrying
3. Industrial Emissions
4. Transport
5. Agriculture
6. Local Measures

Lack of time restricted completion to only two areas, Domestic Combustion and Transport. However, these two areas were sufficient to show how the model could be completed.

The levels in this model are potential policy measures designed to lower concentration of PM_{10}. The policy measures were taken from the supporting studies. The draft model included these measures in cumulative areas, assuming that they were discrete projects but that any, all or none of the options in each area could be undertaken.

When discussing Domestic combustion, the group realised that the options for changing to smokeless fuels and changing to enclosed grates would be better considered as one option. Also, households changing to gas would not also change to smokeless fuels. The first two options were merged to create the option "change to smokeless fuels" and the area was changed to be mutually exclusive.

When considering the options for Transport, the group wanted to recognise that particle traps and retro-fitting would be more effective if sulphur-free diesel had also been introduced. This was therefore fixed as the first option in the Transport area.

The resulting model is shown in Figure 12.

	Domestic Combustion	Quarrying	Industrial Emissions	Transport	Agriculture	Local Measures
8			Non-Ferrous Metals			
7		Improved Road Design	Lime Production			
6		Chemical Suppressants	Cement Production			
5		Enclose Systems	Iron and Steel Industries			
4		Water Sprays	Petroleum Refineries	Retro-Fit Older Vehicles		
3	Convert to Gas	Speed Restrictions	Public and Commercial Boilers	Particulate Traps		
2	Change to Smokeless Fuels	Good Practice Measures	Plant Boilers	Early Sulphur-Free Diesel	?????	?????
1	Do Nothing	Do Nothing	Do Nothing	Do Nothing	Do Nothing	Do Nothing

Figure 12: The model structure for the air quality measures model.

Here, level 1 represents the possibility of enforcing no measures in each area. Even if this is not realistic, it is needed in a model consisting not of mutually exclusive options, as in the air quality policy model, but of cumulative options, so that the added value of an option can be calculated as a difference from the previous level. Any option appearing at level 1 is a given, the starting point for Equity to construct an efficient frontier, so providing a 'Do Nothing' option ensures a level playing field for all six areas.

Although each of the measures is given a brief name, a longer definition was developed in the workshop to ensure clarity. For example the 'Smokeless Fuels' measure means 'Change fuel used domestically to be smokeless. This includes use of closed grates.' And 'Gas' means 'Convert homes to gas heating'.

Finding 13: The MCDA portfolio model can improve the efficiency of budget allocations. The key idea behind this model is that every option is characterised by its benefit-to-cost ratio, a value-for-money index that is acknowledged as the appropriate basis on which to allocate limited resources by corporate finance textbooks[30] as well as *The Green Book* itself[31] This index provides a way to deal with opportunity costs, an important consideration in CBA for deciding what constitutes a cost. The portfolio MCDA model, by including in the analysis other areas in which investments might be made, explicitly takes account of opportunity costs. By so doing, the approach recognises that individually optimal decisions are rarely collectively optimal. For example, the best air quality policy for each of the three regions, shown in Figure 11, considered separately would most likely not be best for the nation as a whole. Investing £X in London, £Y in Scotland and £Z for the rest of the country might be the best thing to do for each area by itself, but the total budget of £X+Y+Z might be distributed differently by recognising that an incremental investment in one area would not achieve the same total benefit as well as the same incremental spend elsewhere. MCDA portfolio models make the trade-offs between investment areas explicit, thereby highlighting the best overall use of the limited resource. This can contribute to decision making that is integrated across different budget areas.

[30] See, for example, Brealey, R. A., Myers, S. C., and Marcus, A. J. (1995) *Fundamentals of Corporate Finance*, New York: McGraw-Hill, p. 150.
[31] *The Green Book*, ibid., paragraph 6.4 and Box 19, page 37.

6.3 Objectives and criteria

Costs

For both models, the cost criteria were:

1. Road Transport Cost
2. Domestic Cost
3. Industry Cost

In each case, these criteria represent the total cost to be borne by all involved parties including (but not limited to) central government, local government, EA, industry and the general public.

For the sake of simplicity in the air quality policy workshop, the total costs were all recorded against one criterion, Industry Cost. Cost figures are taken from the economic analysis (page 66, table 3.19). The group recognised that costs were uncertain and represented an indicative selection of policy measures.

For the air quality measures workshop, the Domestic Combustion and Road Transport costs were, of course, limited to domestic and road transport costs, respectively.

Benefits—policy model

For the air quality policy model the group considered many benefit criteria and agreed the following four:

1. Daily Deaths – a measure of how annual mean concentration effects the death rate
2. Days of Poor Air – a measure of variance from the annual mean to indicate strength of effect on acute mortality and morbidity
3. Life Years Gained – a measure of the long term mortality effects of reducing PM_{10} concentrations
4. Climate Change – a measure of the effect on global warming of PM_{10} measures that also have an impact on CO_2 emissions.

Preference values on the first three criteria were expressed as an inverse function of the input data. This is because the act of reducing particulate concentrations improves health, e.g., fewer deaths means an increase in preference value. The Climate Change criterion does not use an inverse scale as the group agreed that reducing particulate concentrations would generally increase global warming.

It was noted that the climate change estimates were very uncertain but that this would not materially affect the model after weighting.

Benefits—measures model

The group considered many benefit criteria and agreed on the following five:

1. Concentration: Reduction of particle concentration, expressed as a population weighted annual mean by 2010.
2. Hot Spots: Effect on PM_{10} hot spots.
3. Other Effects: An amalgam of minor benefit criteria that may not be relevant for all measures or sources.
4. Government Loss: Loss of revenue, recognising that some measures may lower tax revenues.
5. Progressivity: A measure of the distributional impacts of the policy measure.

Although these five were eventually agreed, there was an intense debate over the benefit criteria. The first area of concern was the number of criteria to be included in the model. Through planning for the workshop, participants had been encouraged to consider all the criteria that could be included that are unsuitable or unusable in traditional analysis methods. A brainstorming session elicited 21 possible criteria. The group expectation was that all these new criteria could be included in the model. There was a degree of confusion when the lead facilitator explained that Equity MCDA models work best with about five to six, at most, carefully chosen benefit criteria. The reason for this is that as more criteria are added, their correlation with each other adds less and less discrimination between the options.

Several resolutions to this debate are possible:

1. Minor criteria were aggregated into the criterion "Other Effects" to save time constructing the model.
2. It was acknowledged that there is no theoretical limit to the number of criteria that can be used in a model but that only a few will materially affect outcomes.
3. MCDA is a long process with three major phases that demand different views of benefit criteria. The first phase is a brainstorming phase where many criteria are identified. Secondly, criteria are filtered for materiality to create and analyse the model. (As above, any number of criteria can theoretically be used.) Lastly, communication with interested parties requires proof that everybody's criteria have been taken into account. Thus, more criteria than is requisitely required might need to be shown.
4. In many cases, criteria are sufficiently related that they can be clustered under higher-level criteria. As a result, two different options may gain similar scores on the high-level criterion but for different reasons.

Finding 14: Debates about what criteria to include in an MCDA can be informed by desirable properties discussed in the MCDA literature[32]. Here is a summary of those properties. The criteria should include all those that are important, yet only as many as are requisite. Redundant criteria should be amalgamated, and double-counting avoided. Each criterion should embody implicit value judgements, making clear what is meant by more or less preferred. Measurements or judgements must be capable of being made about the consequences of the options on the criteria. Definitions of criteria should be unambiguous, with reference to concrete facts. If consequences occur in different time periods, criteria should be included that reflect the value of impacts distributed in time, either with discrete criteria whose spans of time are defined, or with a single criterion expressing the sum of discounted benefits over a defined time period. Finally, if the additive models shown in this report are to be used, the criteria should be mutually preference independent: preference values can be assigned for the options on one criterion without reference to the values on any other criterion, and this should be true for all criteria.

The second area of concern involved the loss of government tax income through incentives to convert to sulphur-free diesel. Projections from the Department of Trade and Industry (DTI) show that tax income will drop during the introduction of sulphur-free diesel. This is due to tax incentives used to encourage people to use the new fuel. The confusion on this issue centred on how this loss of income should be treated in the MCDA model. Traditionally, government departments would treat this as a cost of the policy. This would place the Loss of Revenue criterion in the Costs part of the model.

The MCDA treatment of loss of revenue is as a negative benefit. The argument is that the loss of revenue is a consequence of the policy and therefore not under government control. Whilst a tax incentive can be provided, it is not within government control whether or not the public make use of this provision. In addition, Equity is an input-output model, with inputs reflecting resources that must be committed in order to realise the outputs, the benefits, both positive and negative. This structuring allows Equity to calculate the ratio of the increment in benefits (risk adjusted where appropriate) to the input total costs. This debate was not concluded to the satisfaction of all participants. Some confusion remained but loss of government revenue was eventually treated as a negative benefit.

Finding 15: The distinction between inputs and outputs in the MCDA portfolio model helps to clarify what are costs and what are benefits. Further to Finding 1, decisions, such as policies and measures, are under the control of the decision maker. Decisions require the expenditure of resources, which are, therefore, treated as inputs. The consequences of the policies or measures are only partially, if at all, under the control of the decision maker, even if the consequences can be forecast with certainty. Committing resources can result in undesirable consequences, such as lost tax revenues; these should be seen as negative outputs, treated as dis-benefits. If a policy results in a loss of revenue, that is treated as an output; if it results in a gain in revenue it is still an output. Either way, it would not be set against input costs. If the focus of a CBA is on the *difference* between costs and benefits, rather than the *ratio*, then, lost revenue can be treated as a cost, since the order of adding and subtracting does not matter.

[32] See Keeney (1992), ibid., pp.112-18. Also the *MCA Manual*, ibid., pp. 37-40.

6.4 Scoring

Policies model

Limited time permitted completion of only one area, UK except London and Scotland.

Much discussion attended the group's appraisals of the policy options against the criteria, particularly for benefit criteria with longer-term effects. Participants were surprised to learn that the effects were almost linearly proportional to the annual mean concentration of PM_{10}. This gave rise to identical scores for each of Daily Deaths, Life Years Gained and Climate Change.

In the absence of hard numerical data, the highest concentration of 40µg/m^3 was given a score of 100. The lowest concentration of 20µg/m^3 was given a score of 50 is it was half of 40µg/m^3. All other options were scored proportionally between these two values.

The benefit criterion Days of Poor Air represents the effect of short term peaks in PM_{10} concentration on acute mortality and morbidity. The group felt that the number of days the 50µg/m^3 24 hour mean concentration was exceeded was proportional to the detrimental effect on acute mortality and morbidity; this number was used as an indicator of the health effect.

Costs were available from the economic study separately for the three criteria, but only for the recommended policy. From that starting point, the total cost was calculated, and the group then estimated the total costs for the remaining options based on a rough non-linear curve of cost against annual mean concentration. Experience of the group members shows that the first step in any reduction of particulate matter would be the cheapest and further increments would become steadily more expensive. The total costs are shown under the Industrial Combustion cost criterion.

The agreed input scores for this area are shown in Figure 13.

Input Data

	Costs			Benefits			
Level	RoadTrans	DomComb	IndComb	DailyDeaths	DaysPoorAir	LifeYrsGained	ClimateChange
1 SQ: 50(80/yr), 40an	0	0	0	100	80	100	100
2 50(35/yr), 32an	0	0	120	80	35	80	80
3 23-25 annual mean	0	0	400	60	12	60	60
4 20 annual mean	0	0	900	50	10	50	50
5 50(7/yr), 20an	0	0	1000	50	7	50	50

Figure 13: Input scores for the 'UK except London & Scotland' area in the air quality policy model.

Measures model

At this stage in the modelling, the group was keen to see whether or not the criteria would be closely correlated as with the first workshop. Options were appraised for policy measures in two areas, Domestic Combustion and Transport.

Cost information, and concentration reductions were available from the supporting studies. The other benefit criteria, hot spots, other effects, government loss and progressivity were scored directly as value judgements. The process required participants to agree the most preferred option on a given criterion. This was assigned an arbitrary value of 100. The remaining options were compared to this as a standard, so that, for example, an option judged by the group to contribute half as much value was scored at 50. After several options had been scored, consistency checks were carried out on the sums of options. For example, if options A, B and C were scored at 60, 100 and 40, respectively, the group were asked if A and C together would contribute as much value as 100. This is known as 'balance beam' scoring. This process was only applied to the cumulative options. For mutually exclusive option, differences in scores were compared to each other for consistency.[33]

At this point, the group agreed that the structure of the model was substantially different from the structure of the supporting studies. Most participants believed they could find the required data, given some time. However, it was clear that the bottom-up policy measures approach was a new way of regarding the issue.

Assessments for the two completed areas of the model are shown in Figure 14. A subtle distinction is important to understand for interpreting the data shown, particularly as compared to Figure 13. The options in the policy model, above, are mutually exclusive; one and only one policy can be adopted. Therefore, the costs and benefits shown are those associated with the policy options themselves, as would be expected. For the measures model, below, the options in the Domestic area are also mutually exclusive, so their costs and benefits are as shown. For example, the Convert to Gas option in the Domestic matrix shows a cost of 68.1; this is the estimated cost of converting to gas, while 18.5 is the cost of changing to smokeless fuels. Similarly, on the benefits side, the Concentration of 0.02 for Gas is the benefit of that option. However, the Transport options are cumulative; none, some or all of them could be chosen. Although the costs and benefits are shown as cumulative, the actual cost or benefit of any cumulative option is the increment from the level below it. Thus, the actual cost of retro-fitting older vehicles is 760-730=30. The concentration benefit from particulate traps is 0.05-0.01=0.04.

Input Data

	Costs			Benefits				
Level	Transport	Domestic	Industry	Concentration	Hot Spots	Other Effects	Govt Loss	Progressive
1 Do Nothing	0	0.0	0	0.00	0	0	0	100
2 Change to Smokeless Fuels	0	18.5	0	0.01	10	30	0	75
3 Convert to Gas	0	68.1	0	0.02	100	100	0	0

Input Data

	Costs			Benefits				
Level	Transport	Domestic	Industry	Concentration	Hot Spots	Other Effects	Govt Loss	Progressive
1 Do Nothing	0	0.0	0	0.00	0	20	100	0
2 Early Sulphur-Free Diesel	87	0.0	0	0.01	20	0	30	0
3 Particulate Traps	730	0.0	0	0.05	90	85	10	0
4 Retro-Fit Older Vehicles	760	0.0	0	0.05	100	100	0	0

Figure 14: Input scores for two areas, 'Domestic' (upper), and 'Transport' (lower), in the air quality measures model.

Finding 16: MCDA modelling can accommodate 'top-down' portfolios of policies and 'bottom-up' portfolios of concrete courses of action, here, 'measures'. As long as the criteria satisfy the desirable properties mentioned in Finding 14, MCDA portfolio modelling can be applied to high-level policies and strategies, or operational plans and tactics.

33 The theoretical and empirical justification for this approach is the subject of Chapter 7 in von Winterfeldt, D., & Edwards, W. (1986). *Decision Analysis and Behavioral Research*. Cambridge: Cambridge University Press.

6.5 Weighting
Total benefits cannot be calculated until the units of benefit on all the relative preference scales are equated. To do this requires two sets of weights: within-criterion weights which compare the scales from one area to the next on a given criterion, and across-criteria weights which compare the scales given the highest within-criterion weight on each criterion. Thus, the within-criterion weights equate the scales on a given criterion from one area to the next, while the across-criteria weights equate units from one criterion to the next. These weights are not used on the input scales, but rather on the value scales created by a linear transformation on the input scales: least preferred becomes 0, most preferred is assigned 100, with the remaining input values scaled proportionately.

Policy model
As only one area, UK except London and Scotland, was completed for the policy model, no within-criterion weights were necessary. To assess the across-criteria weights, participants were invited to compare the swings in preference from the least to most preferred options on each of the criteria. For example the difference in concentration between the status quo and most stringent policy was 40μg/m^3 minus 20μg/m^3 which is 20μg/m^3. With the dose-response relationship of +0.75% per 10μg/m^3, the 20μg/m^3 reduction would represent a 1.5% reduction in mortality. The biggest difference in days of poor air is 80 minus 7, or 73. The group were asked which difference matters more, a 1.5% reduction in mortality or a reduction of 73 poor air days. The group agreed it was the mortality reduction. The swing on this criterion was than compared with that of life years gained, with the latter judged as the larger swing. Finally, the swing on climate change was judged to be the least important. Thus, a weight of 100 was assigned to the largest swing, life years gained, and the group judged the sizes of the swings on the remaining criteria, with the following results:

Daily Deaths	Days Poor Air	Life Years Gained	Climate Change
75	50	100	5

Measures model
Because two areas, Domestic and Transport, were completed, it was necessary for the group to assess within-criterion weights for each criterion. As the first step in weighting, the group considered the effect of measures on the two populated areas. Given that the two areas had been scored using the same units (μg/m^3), they could be weighted automatically. Equity assigns weights to the 0-100 value scales for the concentration criterion based on the comparative lengths of the input scales. A clarification question was asked of the group "Does a unit reduction in concentration have the same importance if gained from domestic combustion as it would if gained from transport?" The group answered they were indifferent and therefore automatic weighting was valid. Thus, the 0 to 0.02 swing in reduction for Domestic and the 0 to 0.05 swing in reduction for Transport were given relative weights of 40 and 100.

The other four benefit criteria were not measured in similar units as they were relative judgments; a unit of difference on one scale was not necessarily equivalent to a unit on another. Because of this, the weights were judged by the group. For the criterion Hot Spots, participants were asked the question, "Which is more preferable, the benefit to hot spots from moving to gas in the domestic combustion measures or the benefit from all the transport measures?" The answer was transport so transport was given a weight of 100. Next the group was asked the question, "How beneficial is the effect of moving to gas relative to all the transport measures?" The answer was 40% so domestic combustion was given a weight of 40. Similar questions were asked for all the other criteria, thus completing the elicitation of the within criterion weights.

The group then assessed the across criteria weights. Within-criterion weights of 100 had been assigned to Transport on the first four criteria, and to Domestic on Progressivity. The swings in value on these five criterion scales were compared. The group were asked, "Of all the benefits achievable in all these scales, which criterion shows the largest swing in benefit from 0 to 100?" The answer was Concentration so this was assigned an across-criteria weight of 100. Similar questions about the relativity of benefit were asked and across-criterioa weights were assigned to all the other benefit criteria. The resulting swing weights are shown here:

	Costs			Benefits				
	Transport	Domestic	Industry	Concentration	Hot Spots	Other Effects	Govt Loss	Progressive
Domestic Combustion	-	-	-	40	20	10	0	100
Quarrying	-	-	-	0	0	0	0	0
Industrial Emissions	-	-	-	0	0	0	0	0
Transport	-	-	-	100	100	100	100	0
Agriculture	-	-	-	0	0	0	0	0
Local Measures	-	-	-	0	0	0	0	0
Across Weights	100	100	100	100	20	10	5	10

Done

The zero weight for Government Loss in the Domestic area indicates there is no loss, and similarly for the zero weight for Progressive in the Transport area.

With scoring and weighting finished, the models were now complete.

Finding 17: Weighting in portfolio MCDA is more complex than in simple appraisal MCDA. The added complexity is the result of including several areas, each of which acts as a simple appraisal MCDA, requiring comparisons to be made between the areas. These comparisons take the form of within-criterion weights, which effectively quantify the trade-offs between areas, criterion by criterion. As in the simple appraisal MCDA, swing-weighting is the preferred assessment technique, with the implied trade-offs serving as consistency checks. Throughout the weighting process, participants had to be reminded that it was the swing in value from least to most preferred options they were judging, not the absolute importance of the criterion. This concept is difficult to grasp at first hearing, but eventually it becomes clear. In fact, it is a well-defined task with concrete options serving to anchor the ends of the scale, an easier task than the vague notion of the 'importance' of a criterion.

6.6 Combine weights and scores
6.7 Results
For simplicity, these two steps are combined here. Equity doubly weights the 0-100 value scales and sums the weighted values across the criteria to give a single, total benefit scale for each option. If multiple costs are present, the across-cost weights are applied to give a total cost. For both the policy and measures models, the cost weights are 100 for each criterion, indicating that 100% of each cost is added to obtain the total.

Policy model
Only one area, UK except London and Scotland, was completed. Figure 15 shows the completed input scores and weights. Note that the scores for Daily Deaths, Life Years Gained and Climate change are the same. Recall that this is a result of the judgement that they are all proportional to the annual mean concentration.

Input Data

Level	Costs: RoadTrans	Costs: DomComb	Costs: IndComb	Benefits: DailyDeaths	Benefits: DaysPoorAir	Benefits: LifeYrsGained	Benefits: ClimateChange
1 SQ: 50(80/yr), 40an	0	0	0	100	80	100	100
2 50(35/yr), 32an	0	0	120	80	35	80	80
3 23-25 annual mean	0	0	400	60	12	60	60
4 20 annual mean	0	0	900	50	10	50	50
5 50(7/yr), 20an	0	0	1000	50	7	50	50
Within criteria wts				75	50	100	5
Across criteria wts	100	100	100	100	100	100	100

Figure 15: Input data, scores and weights for the area 'UK except London and Scotland' in the air quality policy model.

The first step to the final results is to normalise all the benefit scales so they extend from 0, representing the least preferred option, to 100, the most preferred. This is shown in Figure 16.

Unweighted Preference Values

Level	Costs: RoadTrans	Costs: DomComb	Costs: IndComb	Benefits: DailyDeaths	Benefits: DaysPoorAir	Benefits: LifeYrsGained	Benefits: ClimateChange
1 SQ: 50(80/yr), 40an	0	0	0	0	0	0	100
2 50(35/yr), 32an	0	0	120	40	62	40	60
3 23-25 annual mean	0	0	400	80	93	80	20
4 20 annual mean	0	0	900	100	96	100	0
5 50(7/yr), 20an	0	0	1000	100	100	100	0
Within criteria wts				75	50	100	5
Across criteria wts	100	100	100	100	100	100	100

Figure 16: Unweighted preference values, the result of normalising the scales shown in Figure 15. Recall that for the first three benefit scores, smaller input numbers are more preferred.

The next step is to doubly weight the benefit preference values, and apply a constant to ensure the total benefits over the whole model don't exceed 1000. This is shown in Figure 17.

Weighted Preference Values

Level	Costs: RoadTrans	Costs: DomComb	Costs: IndComb	Costs: Total	Benefits: DailyDeaths	Benefits: DaysPoorAir	Benefits: LifeYrsGained	Benefits: ClimateChange	Benefits: Total	Ratio
1 SQ: 50(80/yr), 40an	0	0	0	**0**	0	0	0	22	**22**	
2 50(35/yr), 32an	0	0	120	**120**	130	134	174	13	**451**	3.58
3 23-25 annual mean	0	0	400	**400**	261	203	348	4	**816**	1.30
4 20 annual mean	0	0	900	**900**	326	208	435	0	**969**	0.31
5 50(7/yr), 20an	0	0	1000	**1000**	326	217	435	0	**978**	0.09

Figure 17: Weighted preference values for the 'UK except London and Scotland' area showing the total costs and total doubly-weighted benefits, along with the Δbenefit-Δcost ratios in the Ratio column.

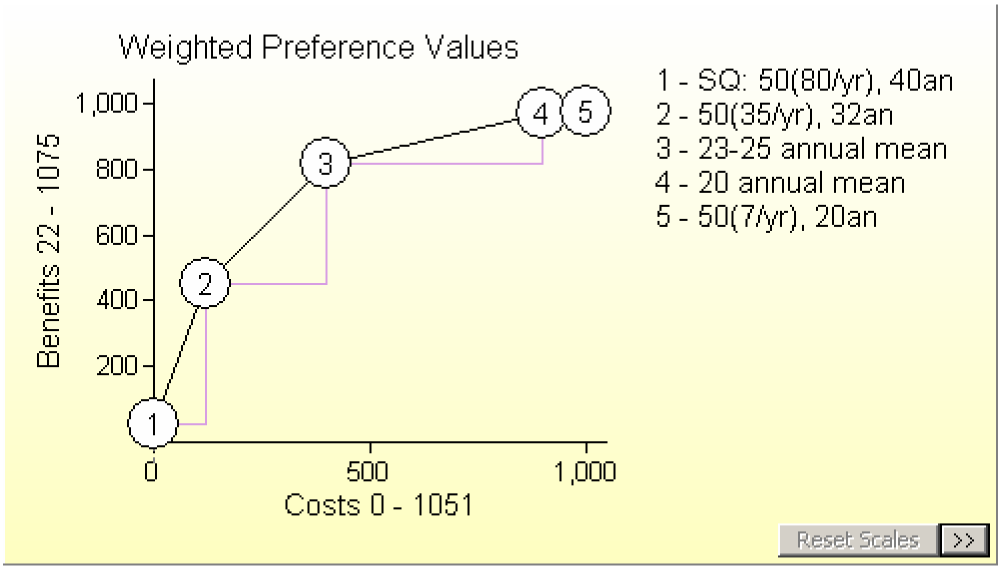


Figure 18: Benefits versus costs for the area 'UK except London and Scotland' in the air quality policy model.

Plotting the total benefits versus the total costs gives the graph shown in Figure 18.

The lightly outlined triangles highlight the increment in benefit that is obtained from the increment in cost as more resource is devoted to the next policy option. The slope of each triangle is an indication of additional value for additional resource. Here, it is clear that the 50(35/yr), 32an, option gives best value for money over and above the status quo of 50(80/yr), 40an. Remember that costs and benefits are in different units, and at this stage no attempt had been made to find the cost equivalent of a unit of benefit. Figure 17 shows the actual Δbenefit-Δcost ratios in the right hand column. For example, the first ratio, 3.5, is obtained by dividing the benefit difference, 451 – 22 = 429 by the corresponding cost difference, 120 – 0 = 120, giving 429/120 = 3.58.

The group agreed that the steps to option 3 looked like good value for money, the step to option 4 was marginal and the step to option 5 seemed to add very little relative benefit.

Measures model

Figures 19 and 20 show the resulting input matrices, weighted preferences and benefit/cost curves for the two areas of the model that were scored and weighted, Domestic Combustion and Transport.

The graph for Domestic Combustion shows a better benefit-to-cost ratio for smokeless fuels than for gas, by 5.18 to 1.78, nearly 3 to 1.

The graph for Transport appears to show nearly a straight line, but reference to the ratios in the weighted preference values display indicates that while the benefit-to-cost ratio for early sulphur-free diesel is somewhat better at 1.12 than particulate traps at 0.83, retro-fitting older vehicles, 0.53, is less than half the benefit-to-cost ratio of sulphur-free diesel. Note, too the relatively large cost of particulate traps.

Finding 18: Displaying the Δbenefit/Δcost triangles for each option makes clear the extent of extra benefit expected from an extra investment. Although the same broad picture emerged here as from the CBA cost curves, the graphical display provides a transparent indication of the benefit-to-cost ratio.

Input Data

Level	Costs: Transport	Costs: Domestic	Costs: Industry	Benefits: Concentration	Benefits: Hot Spots	Benefits: Other Effects	Benefits: Govt Loss	Benefits: Progressive
1 Do Nothing	0	0.0	0	0.00	0	0	0	100
2 Change to Smokeless Fuels	0	18.5	0	0.01	10	30	0	75
3 Convert to Gas	0	68.1	0	0.02	100	100	0	0
Within criteria wts				40	20	10	0	100
Across criteria wts	100	100	100	100	20	10	5	10

Weighted Preference Values

Level	Costs: Transport	Costs: Domestic	Costs: Industry	Costs: Total	Benefits: Concentration	Benefits: Hot Spots	Benefits: Other Effects	Benefits: Govt Loss	Benefits: Progressive	Benefits: Total	Ratio
1 Do Nothing	0	0.0	0	**0**	0	0	0	0	53	**53**	
2 Change to Smokeless Fuels	0	18.5	0	**18**	105	2	2	0	39	**148**	5.18
3 Convert to Gas	0	68.1	0	**68**	211	21	5	0	0	**237**	1.78

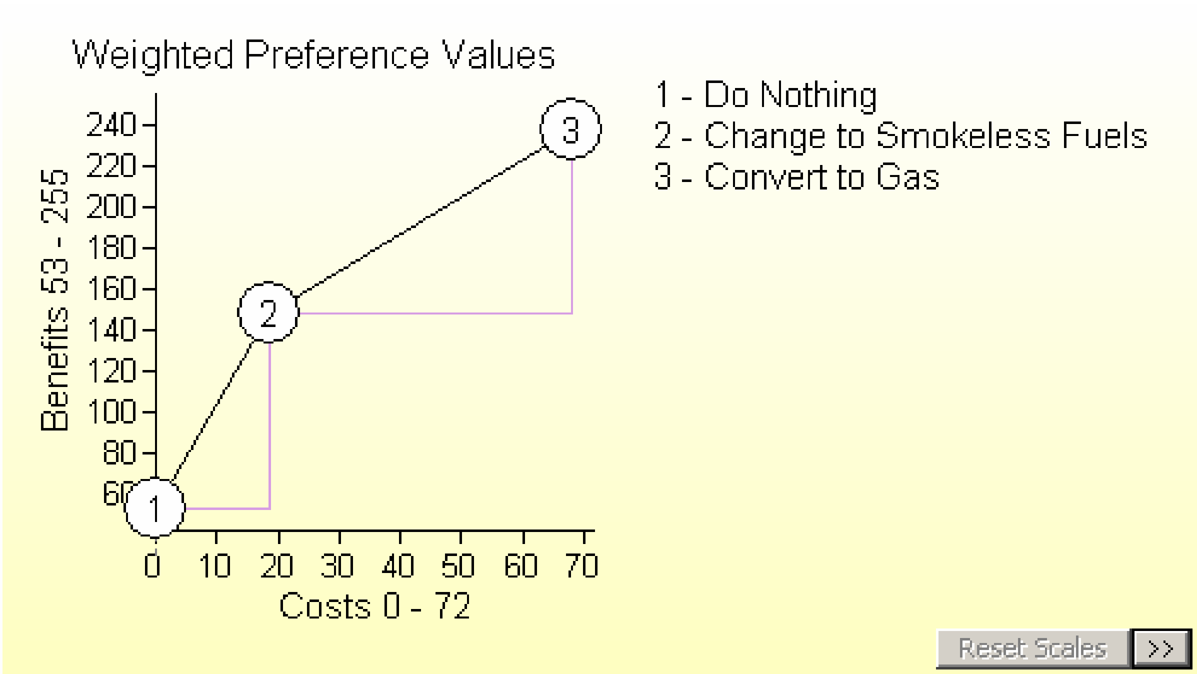


Figure 19: Input scores and weights, weighted preference values and graph of total benefits versus total costs for the Domestic Combustion area in the air quality measures model.

Because the units of benefit on these two graphs have been equated, it is possible to combine the five triangles into one graph. It is shown in Figure 21. With three measures options for Domestic Combustion and four for Transport, there are 12 possible combinations of measures for just these two areas. All those 12 combinations appear somewhere within the green (shaded) area of the curve, but only 5 appear on the efficient frontier. The five points represent the best portfolio of options for a given cost, where any point upwards and to the right includes in the portfolio all cumulative options to the left. The Order of Priority shows the identity of the successive points.

Input Data

Level	Costs: Transport	Costs: Domestic	Costs: Industry	Benefits: Concentration	Benefits: Hot Spots	Benefits: Other Effects	Benefits: Govt Loss	Benefits: Progressive
1 Do Nothing	0	0.0	0	0.00	0	20	100	0
2 Early Sulphur-Free Diesel	87	0.0	0	0.01	20	0	30	0
3 Particulate Traps	730	0.0	0	0.05	90	85	10	0
4 Retro-Fit Older Vehicles	760	0.0	0	0.05	100	100	0	0
Within criteria wts				100	100	100	100	0
Across criteria wts	100	100	100	100	20	10	5	10

Weighted Preference Values

Level	Costs: Transport	Costs: Domestic	Costs: Industry	Costs: Total	Benefits: Concentration	Benefits: Hot Spots	Benefits: Other Effects	Benefits: Govt Loss	Benefits: Progressive	Benefits: Total	Ratio
1 Do Nothing	0	0.0	0	**0**	0	0	11	26	0	**37**	1.12
2 Early Sulphur-Free Diesel	87	0.0	0	**87**	105	21	0	8	0	**134**	0.83
3 Particulate Traps	730	0.0	0	**730**	526	95	45	3	0	**668**	0.53
4 Retro-Fit Older Vehicles	760	0.0	0	**760**	526	105	53	0	0	**684**	

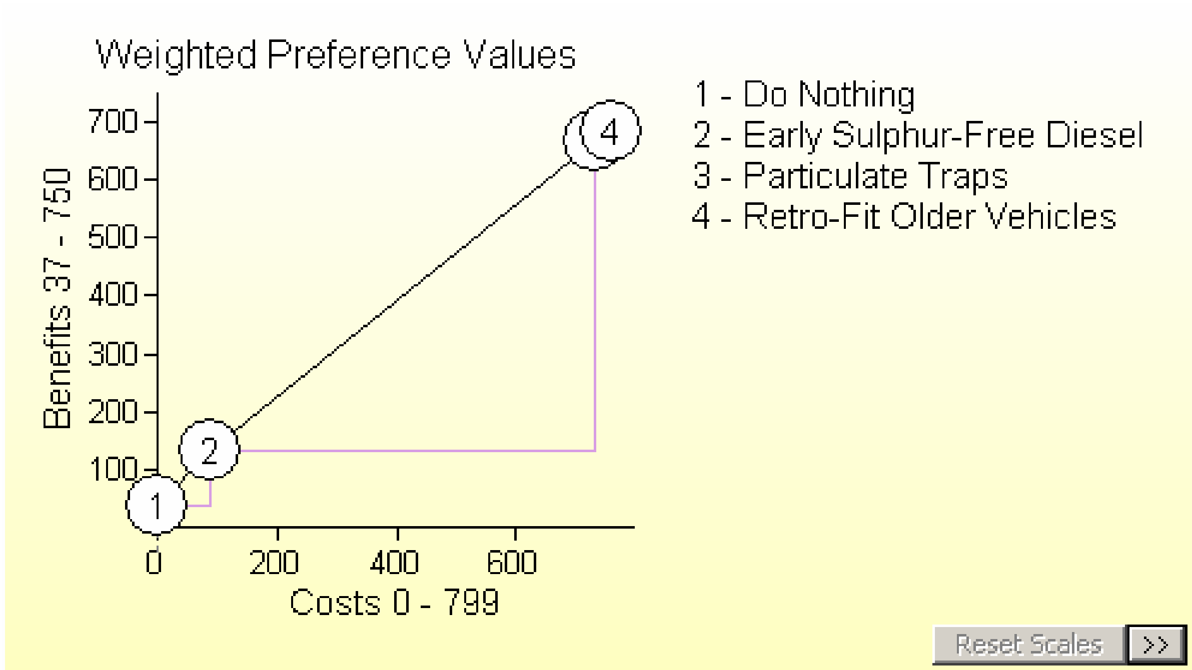
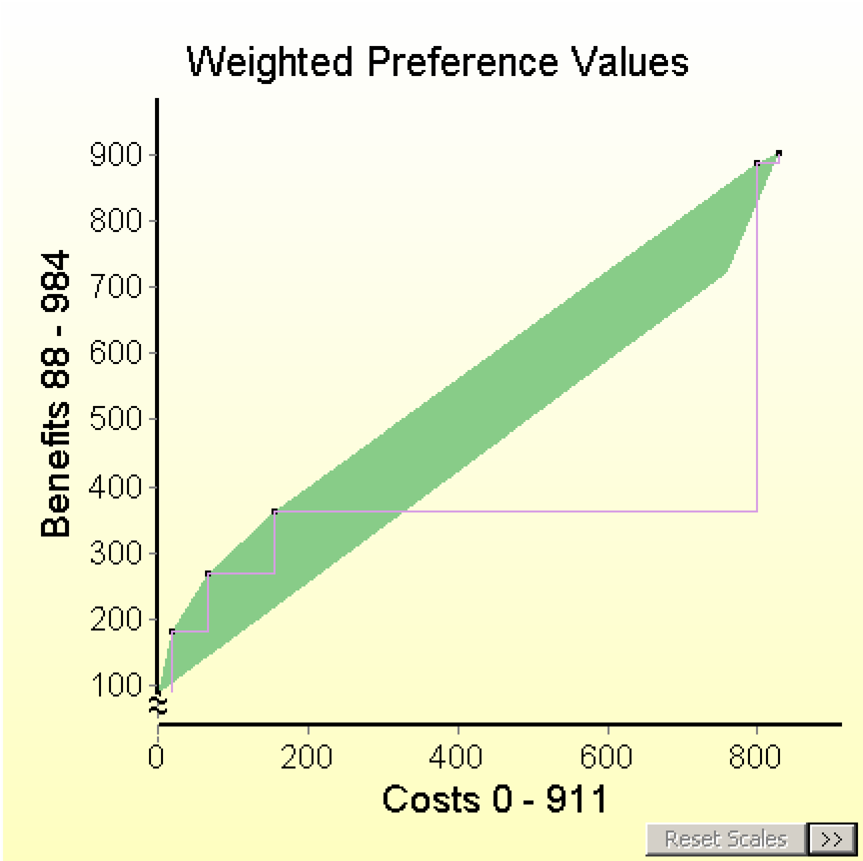


Figure 20: Input scores and weights, weighted preference values and graph of total benefits versus total costs for the Transport area in the air quality measures model.

Note that the first two points are both from Domestic Combustion, followed by the three Transport options. This shows that relatively speaking, the domestic combustion options are better value for money than the transport measures. The curve also shows the very substantial relative cost of particulate traps compared to the other measures.



Order Of Priority

	Area		Level		Costs Inc	Costs Cum	Benefits Inc	Benefits Cum
0	1	Domestic Combustion	1	Do Nothing	0	0	52	52
0	2	Quarrying	1	Do Nothing	0	0	0	52
0	3	Industrial Emissions	1	Do Nothing	0	0	0	52
0	4	Transport	1	Do Nothing	0	0	36	88
0	5	Agriculture	1	Do Nothing	0	0	0	88
0	6	Local Measures	1	Do Nothing	0	0	0	88
1	1	Domestic Combustion	2	Change to Smokeless Fuels	18	18	94	181
2	1	Domestic Combustion	3	Convert to Gas	50	68	87	268
3	4	Transport	2	Early Sulphur-Free Diesel	87	155	95	363
4	4	Transport	3	Particulate Traps	643	798	523	887
5	4	Transport	4	Retro-Fit Older Vehicles	30	828	15	902

Figure 21: The efficient frontier comprising options from the Domestic Combustion and Transport areas for the measures air quality model. The Order of Priority shows the successive points on the curve, from the lower left starting point, Do Nothing in all areas.

If time had permitted the group to finish the model, many more points would, of course, have appeared on the efficient frontier, and an interweaving of options from different areas would have occurred. This would enable policy makers to see the relative priorities of all the measures. It would also make clear that if the best measure from each area were selected, that particular portfolio of measures would not have been the best use of the total resource. Here, for example, selecting Change to Smokeless Fuels and Early Sulphur Diesel would have cost 105 and given total benefits of 189, a ratio of 1.80. A better portfolio would be to convert to gas, for a total cost of 68 and total benefits of 268, a ratio of 3.94.

Finding 19: Portfolio MCDA modelling overcomes the inefficient use of resources that results from 'silo' decisions. Setting a common air quality standard for all parts of the UK would result in the inefficient use of resources because a unit of cost would not achieve the same benefits in each geographical sector. This recognises the existence of inequality of opportunities, which require investments in the different areas that equate the Δbenefit/Δcost ratios among the areas, not the benefits themselves, or the differences between benefits and costs. In doing this, the 'best' portfolio overcomes the 'commons dilemma'[34]: a common resource used to the individual best advantage of each user is not the collectively best use of the resource. Experience with portfolio MCDA modelling shows that the collectively best portfolio provides on average 30% more benefit than the collection of separate 'silo' uses. Thus, integrated decision making, the topic of Finding 13, can realise substantially more benefits at no increase in resource.

6.8 Sensitivity analyses

In modelling portfolios, sensitivity analysis becomes a matter of seeing whether different scores or weights affect the Order of Priority of the options. One way to show the effects is to see the impact on the benefit/cost ratios.

Policy model

During the workshop, there was much discussion within the group about assigned values, both scores and weights. Specific items were picked out for sensitivity analysis:

1. The cost of reducing PM_{10} concentration annual mean from 32μg/m^3 to 23-25μg/m^3.
2. The weighting of Days of Poor Air.
3. The weighting of Climate Change.

The input data were changed to explore the sensitivity of these parts of the model. Results are shown numerically in terms of the change in benefit/cost ratios in Table 5.

It can be seen from the ratios that the three areas of dispute within the group did not substantially change the results. The steps to Option 3 were still steep with a gradient of 1 or greater. The steps beyond Option 3 were shallow with gradients far less than 1. The relative ordering of the options within their areas remained unchanged. The group agreed that the disputed model elements were not sensitive to these different views.

[34] Hardin, G. (1968). The tragedy of the commons. *Science, 162*, 1243-1248.

Table 5: Changes is benefit/cost ratios as costs and weights are varied over a range of values.

	B/C ratios of 0ptions for different costs of Option 3		
Option	**£300m**	**£400m (base case)**	**£500m**
2. 32μg/m^3 annual mean	3.58	3.58	3.58
3. 23-25μg/m^3 annual mean	2.02	1.30	0.96
4. 20μg/m^3 annual mean	0.26	0.31	0.38
5. 20μg/m^3 annual mean	0.09	0.09	0.09

	B/C ratios of 0ptions for different input weights for Poor Air		
Option	**25**	**50 (base case)**	**75**
2. 32μg/m^3 annual mean	3.39	3.58	3.73
3. 23-25μg/m^3 annual mean	1.32	1.30	1.28
4. 20μg/m^3 annual mean	0.34	0.31	0.28
5. 20μg/m^3 annual mean	0.05	0.09	0.12

	B/C ratios of 0ptions for different input weights for Climate		
Option	**5 (base case)**	**15**	**25**
2. 32μg/m^3 annual mean	3.58	3.29	3.03
3. 23-25μg/m^3 annual mean	1.30	1.19	1.08
4. 20μg/m^3 annual mean	0.31	0.28	0.25
5. 20μg/m^3 annual mean	0.09	0.09	0.08

Measures model
Time did not permit the group to carry out any sensitivity analyses. However, one is reported here to show the possible effects on the Order of Priority. Suppose someone argues that the unknown effects on loss of government income is unfavourably and unrealistically biasing the priorities against the Transport options. Reducing that weight to zero turns out to have no impact on the Order of Priority. Then suppose it is argued that increasing the weight on Hot Spots would favour the Transport options. Doubling the across-criterion weight from 20 to 40 again has no impact on the Order of Priority. Thus, over a substantial range of changes, the Transport options remain lower priority than the Domestic Combustion options.

This failure to see any changes in either model is partly caused by the incompleteness of the models. But it mainly is the result of substantial relative difference in the benefit-to-cost ratios. It is typical of MCDA models, as we said before, that their results can tolerate large ranges of differences in input scores and weights. Thus, precision in those input values is not required.

Finding 20: Sensitivity analysis plays a crucial role in MCDA, as it does in CBA. In both appraisal MCDA and portfolio MCDA the substantial insensitivity of overall results to changes in scores and weights is the result of the structure of the models. In appraisal MCDA, the greater the statistical correlation between the criteria (which can occur even if they are mutually preference independent), the less sensitive are the results to the weights. If some criteria are negatively correlated with others, then the weights will matter. Much the same applies to portfolio MCDA, with the added observation that as costs and benefits are often correlated (more cost leads to more benefits), it takes considerable changes to affect the benefit/cost ratios.

6.9 Discussion

The PM_{10} policy model showed how the structure of a portfolio model could be realised, even though only one area was completed. The PM_{10} measures model began to realise the full capability of the portfolio approach. The portfolio model was shown to be a collection of separate appraisal MCDA models that are connected through a weighting system that makes explicit the trade-offs among the areas of the model and among the criteria. Thus, the model recognises the inequalities of opportunity from one area to the next, and that some criteria are relatively more important than others.

Portfolios were shown to be built up by considering the increment in total benefit obtained from an additional increment of financial investment. The denominator of this ratio of Δbenefit/Δcost is in monetary terms, while the numerator is in units of value that have been equated across all areas and criteria. It is this ratio that is the proper basis for decision making when budgets are limited, and it is crucial in constructing portfolios of options.

Graphical display of benefits versus costs demonstrably helped participants in the decision conferences to understand the relative merits of the options under consideration. Although no such displays were evident in the source materials on which much of the MCDA analyses were based, perhaps this is another argument for the use of MCDA in conjunction with CBA. CBA can provide inputs to MCDA which can then be combined with other non-monetisable criteria, and displayed in ways that deepen insight into the issues. Particularly when the modelling is carried out in an open forum of experts, sensitivity analyses can then be used to explore the consequences of informed differences of opinion, enabling a group to arrive at agreement about the way forward even though they may disagree about specific details which were shown not to affect the overall results.

We are mindful of the difficulty some economists, in particular, experience when first confronted with MCDA, which seems somehow less objective and less able to build the confidence of decision makers. We think this project demonstrates the opposite. When a group of informed experts collectively test their differences of opinion in a multi-criteria model, then revise and further test the model, they are more likely to reach an overall agreement than they would without the benefit of the MCDA model. A decision maker who knows that a group of specialists agree, having tested alternative perspectives, can feel confident in the support of the recommending group. While the limited workshop time in this project was insufficient to create complete alignment of the participants, our experience shows that with only slightly more time to complete and fully explore a model, committed alignment can result.

We conducted this work against a background theory about when enough modelling is enough. The theory of requisite decision models[35] proposes that a model should be just sufficient in form and content to resolve the issues at hand. That, and no more. During this project we were given documents totalling nearly 600 pages, but the actual input used in the MCDA models came from a small fraction of that total. And still, key data were missing. In particular, the single illustrative scenario presented in The Air Quality Strategy focussed on a single option, with the result that full data for the other options were not available. It is tempting for an organisation to devote considerable resources to gathering data, but it is typical that only a fraction of that data is needed for informed decision making. To keep the search for data requisite, an MCDA conducted at the start of a project, mainly using expert judgement as inputs, with many sensitivity analyses conducted to see which data are crucial, can narrow the subsequent search for data to only information that matters to the decisions. Once the data have been collected, another MCDA analysis, perhaps updating the original one, can then bring the problem into sharper focus and lead the specialist group to agree the way forward. But as we said earlier, this is a very different model from the five-stage model currently employed in the Civil Service: research, modelling, consultation, revision, final recommendations. The MCDA approach is very iterative; the eight steps shown in Figure 1 may be cycled through several times before a requisite model is obtained.

In addition, sensitivity analysis would help to discover whether the wide range of uncertainty that attended the CBA analysis on many of the effects actually matter to the prioritisation of the options. Those that do could then be represented with probability distributions and integrated into the MCDA, which can accommodate uncertainty in several ways.

[35] Phillips, L. D. (1984). A theory of requisite decision models. *Acta Psychologica, 56*, 29-48.

6.10 Summary

The two air quality MCDA models examined portfolios of options, the first, portfolios of policy options for three geographical divisions of the UK, and the second, portfolios of specific measures to be taken by six sectors of polluters.

For the policy options model, slightly different sets of options were proposed for each of three geographical areas: UK except London and Scotland, London, and Scotland. Options were appraised for only one area, UK except London and Scotland, but sufficient information was available to appraise the five policy options against four criteria, daily deaths, days of poor air, life years gained and climate change. Compared to the current policy, 24-hour mean of 50 $\mu g/m^3$ not to be exceeded more than 80 times per year with an annual mean of $40\mu g/m^3$ by 2005, the model showed reasonably high value-for-money ratios for two policies: a 24-hour mean of 50 $\mu g/m^3$ not to be exceeded more than 35 times per year with an annual mean of $32ug/m^3$, and an annual mean of 23-25 $\mu g/m^3$. More stringent policies that reduce the permitted mean or the number of days exceeding the standard were much less cost-efficient.

The measures model proposed different sets of measures that could be taken by six sources of pollution: domestic combustion, quarrying, industrial emissions, transport, agriculture and local measures. Options for domestic combustion and transport were appraised for their costs and benefits against five benefit criteria: reduction in concentration of particles, local effects of policy measures, peripheral criteria including non-health effects, loss of government income through incentives and the extent to which the measure is socially progressive or regressive.

The overall prioritisation of the two domestic combustion measures and three transport measures showed that the domestic combustion measures were higher priority, on a value-for-money basis, than the transport measures. The model demonstrated how relative judgements between areas and criteria can produce a common unit of value across the areas and criteria, enabling portfolios of best options to be constructed.

Eight findings relevant to the comparison of MCDA with CBA emerged from this part of the study:

1. The MCDA portfolio model can provide the basis for 'joined up' decision making.
2. Debates about what criteria to include in an MCDA can be informed by desirable properties discussed in the MCDA literature.
3. The distinction between inputs and outputs in the MCDA portfolio model helps to clarify what is a cost and what is a benefit.
4. MCDA modelling can accommodate 'top-down' portfolios of policies and 'bottom-up' portfolios of concrete courses of action, here, 'measures'.
5. Weighting in portfolio MCDA is more complex than in simple appraisal MCDA.
6. Displaying the Δbenefit/Δcost triangles for each option makes clear the extent of extra benefit expected from an extra investment.
7. Portfolio MCDA modelling overcomes the inefficient use of resources that results from 'silo' decisions.
8. Sensitivity analysis plays a crucial role in MCDA, as it does in CBA.

7. Conclusions

While this study demonstrates the application of portfolio MCDA to appraising options and to the creation, either 'top-down' or 'bottom-up', of portfolios of options, the reader may still be left with questions about how the approach would be implemented in a specific government application, and how it would be different from a CBA. We believe there is scope for further exploration not only of the common ground between the approaches, but also of new ways for supporting government decision.

The best way to move forward is to work together on live projects. We found it difficult to discover the information we needed in those 600 pages, and it was often not in a readily digestible form, or in a form that could be used directly in an MCDA. We were frustrated by the lack of information on all the options; instead, information on the 'illustrative example' was given in great depth, certainly more depth than was need for the MCDA models. That is not to say that in-depth information may not be required by an MCDA; it often is, but usually 'in-depth' means more detail, and usually that detail simply isn't requisite for the decision. Thus, we would find considerable value as members of a team working on an important and topical issue, helping the team to construct a requisite model that would inform the decision maker, and guiding the search for better information. Civil servants are themselves limited resources in that infinite time is not available to explore every issue. It seems to us that the requisite use of human resources is every bit as important as the requisiteness of the models these specialists create.

Where air quality is concerned, the importance of applying MCDA has already been recognised by the Ad Hoc Group on the Economic Appraisal of the Health Effects of Air Pollution. That group's report[36] made the following observation:

"In reality, a need for a consideration of multiple factors is probably inevitable when developing policies to reduce risks to public health since rather few decisions involve a straight trade-off between health risks and money (and *nothing* else). This suggests that any form of political decision-making can be seen as a form of multi-criteria analysis (just as it can also be seen as an informal form of cost-benefit analysis since it will involve consideration of advantages and disadvantages). In everyday life, as we have already noted, people routinely trade off personal safety against a variety of other *non-monetary* criteria such as enjoyment, or time or convenience. In doing so, they "solve" a multiple criteria decision problem – without going through the process of putting a monetary value on safety, another on enjoyment, time or convenience and then comparing the answers." (paragraph 3.44, page 39; italics theirs)

It is important to recognise that MCDA was not developed from a base of welfare economics. Therefore, it has not derived procedures for valuing monetisable benefits and costs that are specially geared to the public interest. We are convinced that procedures developed within the CBA community could provide valuable inputs to monetary criteria in MCDA analyses.

It is our view that the power of CBA to arrive at informed monetary values for effects that can be monetised, combined with the flexibility and scope of MCDA to handle non-monetary values and trade-offs between conflicting criteria, would together provide a more fully-reasoned, supported and transparent analysis to aid government decision makers than either could alone.

Acknowledgements

We offer our particular thanks to Michael Spackman for his invaluable help in clarifying aspects of the theory and practice of cost-benefit analysis, particularly as it is used within the UK Government. We are also grateful to the DEFRA project team whose many suggestions helped us to improve this report, but we take full responsibility for any remaining omissions and errors.

[36] Department of Health (1999). *Economic Appraisal of the Health Effects of Air Pollution*. London: The Stationery Office.

8. Glossary of terms

Across-criteria weights. Numbers representing the equivalence of a unit of value on one criterion scale with the unit on a different criterion scale. In Equity, these are assessed after the within-criterion weights are determined. Then, those scales assigned within-criterion weights of 100 are compared using the 'swing weight' method (see Within-criterion weights). Across criteria weights reflect the relative importance of the criteria. In Hiview, only across-criteria weights are assessed.

Budget category. An area for allocating resources. In Equity, these are separate areas within which options are evaluated against cost and benefit criteria.

Consequences. The end results from making a choice and observing or experiencing the outcomes.

Cost-effectiveness. In Equity, cost-effectiveness of a project is judged by the ratio of increment of benefit to increment of cost, an index of 'value-for-money'.

Criterion. A standard against which projects are evaluated in terms of preference value.

Goal. A measurable objective. A goal is stated with sufficient detail that it would be possible, at least in principle, to measure at the appropriate time in the future whether the objective has been achieved. For example, 'reduce the concentration of PM_{10}' would be an objective, whereas 'reduce the 24-hour mean concentration of PM_{10}, as measured in a specified location, to 50 $\mu g/m^3$ not to be exceeded more than 80 times per year and an annual mean of $40\mu g/m^3$ by 2005' would be a goal.

Objective. A desired end state at some time in the future.

Option. A description of a possible choice that could be made by the decision maker. The option could be a policy, a strategy, a system, an abatement measure, or anything that is completely under the control of the decision maker. A final decision is made at the point when a decision maker approves the resource that is needed to effect the decision.

Order of Priority. The projects from all budget categories ordered by the ratio of their associated increment of benefit to increment of cost.

Preference scale. A scale representing relative strength of preference, where 100 and 0 are associated with more and less preferred reference points. The two points can be 'Good' and 'Bad, 'Maximum Feasible' and 'Minimum Acceptable', or any other descriptions that define a range of possible positions. This includes the common practice of defining the points in terms of the options being considered, with the most preferred option for a given criterion assigned 100, and the least preferred option a zero. Note that this is a relative scale, so 0 does not mean 'no preference value', any more than 0° Celsius means no temperature.

Trade-offs. A judgment expressed as a numerical weight, representing the extent to which an increment of value on one criterion is equivalent to an increment of value on another criterion.

Value, or Preference Value. The extent to which an option achieves an objective.

Weight: A numerical measure associated with a criterion that reflects how much the two reference points on the criterion (see **Preference Scale**) differ from each other, and how much the decision maker cares about that difference.

Within-criterion weights. Numbers representing the equivalence of a unit of value on one scale with the unit on another scale, for a given criterion. In Equity, these weights reflect the swings in value from one budget category to the next, on a given criterion. The weights are assessed by the 'swing weight' method: judging the swing in preference from the bottom to the top of one scale as compared to the bottom-to-top swing on another scale, where the swings in preference are based on the magnitude of the difference between the least and most preferred levels, and how much the assessor cares about that difference.