

Aversion to health inequalities in healthcare prioritisation: a multicriteria optimisation perspective

Alec Morton^a

^a*Department of Management Science, Strathclyde Business School, University of
Strathclyde, 16 Richmond Street, Glasgow G1 1XQ, United Kingdom. E-mail:
alec.morton@strath.ac.uk*

Abstract

In this paper we discuss the prioritisation of healthcare projects where there is a concern about health inequalities, but the decision maker is reluctant to make explicit quantitative value judgements and the data systems only allow the measurement of health at an aggregate level. Our analysis begins with a standard welfare economic model of healthcare resource allocation. We show how - under the assumption that the healthcare projects under consideration have a small impact on individual health - the problem can be reformulated as one of finding a particular subset of the class of efficient solutions to an implied multicriteria optimisation problem. Algorithms for finding such solutions are readily available, and we demonstrate our approach through a worked example of treatment for clinical depression.

Key words: health inequalities, mathematical programming, multicriteria decision analysis (MCDA)

1. Introduction

How best to take into account health inequalities is an ongoing issue in the theory and practice of health economics. In many developed countries, policy makers and the public seem to care about inequalities in population health (see Morton and Airoldi, 2009, for a discussion of the recent history of health inequality policy in the UK), and demonstrate this concern through both policy action and surveys of social values. This preference has been an important pre-occupation for health economists over the last two decades or so (Williams and Cookson, 2000; McIntyre and Mooney, 2007). However, it is still not clear how concerns about inequality should be captured in practical appraisal procedures, whether at the national level (for example, economic evaluation of technologies, development of national guidelines) or at the local level (prioritisation of spend by health authorities on the ground). A particular difficulty in connecting theory to practice is that health planners are often very reluctant to explicitly state parameters that reflect differences in importance of one section of the population rather than another: it is hard to imagine a Minister of Health announcing that

a QALY accruing to a smoker was to be valued as 80% of a QALY for a non-smoker, for example. Yet most forms of quantitative modelling seem to require such explicit tradeoff statements. A further difficulty in connecting theory to practice is that most health planners work with aggregate data and population averages, and do not have access to information about individual members of the population - indeed for many purposes, information about the distribution of health status in the entire population may simply not be collected, and planners have to rely on extrapolation from surveys or field studies. Yet welfare economic models construct their models of societal value by building upwards from an individual base (see Østerdal, 2005).

In this paper, we consider the question of how to prioritise in the face of incomplete information about values and aggregated information about population health through the lens of Multicriteria Optimisation (MCO). MCO deals with the formulation of and solution procedures for optimisation problems where there are multiple conflicting objective functions which cannot be completely traded off against each other. Solving a MCO involves identifying all solutions that are efficient in the sense that they are optimal with respect to some aggregate objective function within a family of possible functions, rather than optimising a single unique objective function. Where the family of objective function consists of all monotonically increasing functions, this MCO definition of efficiency collapses to the familiar concept of Pareto efficiency. Thus, MCO procedures side-step the problem of explicitly parameterising an objective function - the client for the analysis is presented with a number of "efficient", but possibly very different, solutions, between which they can choose directly. In some cases the identification of efficient solutions may be sufficient to arrive at a choice (e.g. if there a single, or single acceptable, efficient solution); but even in other cases, where multiple efficient solutions exist, there may be advantage in focussing discussions on concrete choices between alternatives, rather than on abstract and polarising questions of how much one values a health benefit to one sort of person rather than another sort of person.

This paper follows in a line of health economic papers that seek to apply mathematical programming approaches in health economics (Birch and Gafni, 1992; 1993; Johannesson and Weinstein, 1993; Stinnett and Paltiel, 1996; Earnshaw and Dennett, 2003; Anand, 2003; Epstein et al., 2007; Cleary et al., 2010). It has long been recognised that the ordering derived from the cost per QALY rule can be interpreted as a heuristic solution to a more general, but implicit, optimisation problem (Weinstein and Zeckhauser, 1973; Dantzig, 1998). These more general mathematical programming formulations model healthcare prioritisation as resource allocation subject to a fixed budget constraint, where the opportunity costs of choosing a particular healthcare investment are explicitly modelled. The mathematical programming framework can take account of issues such as indivisibilities, returns to scale, interactions between alternative investments, and the availability of recourse actions if investment decisions taken under uncertainty do not yield satisfactory results. The paper of Anand (2003) in particular is similar in spirit to the present work in that the author observes the healthcare resource allocation problem evokes conflicting values

and considers how optimal solutions may differ depending on the choice of objective function. However, to the best of our knowledge, MCO has not yet been proposed in a health economics context.

Equity can be captured in various ways in the context of the mono-criterion models already proposed in the literature. For example, a social planner can impose through constraints that a certain amount of resources are devoted, or a certain amount of health is delivered, to some particular group (Stinnett and Paltiel, 1996). Such an approach may - arguably - be appropriate in certain situations (for example, if one population group deserves redress for some previous wrong, then a mathematical program could be constrained so that they are compensated to the extent that they were previously wronged). But it has clear disadvantages, in general. For example, it may be simply impossible to find a solution that meets all constraints. Imposing constraints on lower levels of resource consumption in particular may have perverse effects, where required expenditure limits for one group may only be possible if they are provided with ineffective or indeed harmful treatment.

In this paper we seek to make a number of contributions:

- we introduce MCO concepts and relate these to the welfare economic theory of health as it relates to health inequality;
- we show that the commonly used but *ad hoc* approach of representing equity concerns through constraints recommends solutions that do not satisfy a multicriteria efficiency condition;
- we show through a worked example that MCO is workable technology for healthcare resource allocation, where project selection is modelled through continuous or discrete decision variables.

The structure of the paper is as follows. In Section 2 we outline a basic model of healthcare resource allocation. In Section 3 we present key MCO concepts and demonstrate how they relate to the model of the previous section. In Section 4 we present an example based on prioritising treatments for depression in England, and Section 5 concludes.

2. Our model

2.1. *The (health-related) welfare economic frame*

How should a social planner think about choice between alternative healthcare investments? A common approach in health economics is to assume that such investment decisions are only value-relevant insofar as they have influence on the health of a population captured through a Health-Related Social Welfare Function or HR-SWF. A substantial theoretic literature exists on the underpinning normative theory - see Culyer (1989), Wagstaff (1991), Bleichrodt (1997), Williams and Cookson (2000), Østerdal (2005) and Epstein et al. (2009) - as well as a growing body of empirical research (Dolan, Shaw, Tsuchiya and Williams 2005; Dolan, Edlin and Tsuchiya 2008).

The core idea behind the HR-SWF can be presented as follows. Let $\mathcal{N}=\{1, \dots, i, \dots, N\}$ be the index set for the members of the population. A very general form for the HR-SWF is as follows:

$$\sum_{j \in \mathcal{N}} w_j u(h_j) \tag{1}$$

The variable $h_j \in \mathbb{R}_+$ represents a measure of the lifetime health for person j . The concave increasing function $u : \mathbb{R}_+ \mapsto \mathbb{R}_+$ captures the idea that the more health someone has, the less valuable (to the social planner) a marginal increase is. The scaling factor w_j reflects that the health of certain people may be valued more than other people, because of certain characteristics of these people (for example, some of them may be smokers, or may have dependants). If $w_j = w \forall j$, then this HR-SWF is interpersonally anonymous in the sense that the same health benefit (for example, one QALY) is valued the same when one individual receives it as when another individual at same level of health does.

The HR-SWF presented in (1) is separable in the sense that it is possible to value a health gain to one person without knowing anything about what health gains accrue to other people. This is not as limiting as it might appear: many non-additive HR-SWFs (e.g. the CES function) can be transformed by a monotonically increasing transformation to an additive function, and are thus "strategically equivalent" to an additive function, in the sense that in any optimisation problem, the additive transformed version of the function can replace with original non-additive function without changing the optimal solution. There are however functions – in particular the Rawlsian maximin function – that cannot be transformed thus, and there are situations where one may wish to model interdependencies between persons - for example in the case where there are "caring externalities" (Culyer, 1989)- and so this is not a completely vacuous assumption.

Such welfare models are indispensable for theoretic analyses, but but have limitations for purposes of practical use in an appraisal or resource allocation context. A particularly practical difficulty with this model is that it seems to necessitate measuring the health of every individual in a population and planning based on that individual level data. This is unlikely to be possible. One way round this difficulty is to work with a simpler function. Where health improvements are marginal, this can be justified. In particular, for health improvements $\delta = (\delta_1, \dots, \delta_j, \dots, \delta_N)$ that are "small" in the sense that the first order Taylor series $u(h_j^0 + \delta) = u(h_j^0) + \frac{du(h_j^0)}{dh} \delta_j$ is an adequate approximation of the $u(h_j)$ s, equation (1) can be simplified as:

$$\sum_{j \in \mathcal{N}} w_j u(h_j^0) + \sum_{j \in \mathcal{N}} \frac{du(h_j^0)}{dh} w_j \delta_j \tag{2}$$

Maximising equation (2) by varying δ is thus equivalent to maximising $\sum_{j \in \mathcal{N}} \frac{du(h_j^0)}{dh} w_j \delta_j$

. Thus knowledge of $h^0 = (h_1^0, \dots, h_j^0, \dots, h_N^0)$ is not required directly, except

through the derivatives.

Two arguments that can be used to justify the claim that investments typically have a "small impact on individual health" and hence justify the Taylor series approximation are as follows:

- healthcare consumption occurs disproportionately late in life, and indeed disproportionately shortly before death (see Payne et al., 2007 for a recent review), and so, logically speaking, this consumption can only have a rather marginal impact on lifetime health;
- as is stressed in cost effectiveness textbooks (e.g. Drummond et al., 2005), healthcare investments decisions will normally be incremental (e.g. the decision to upgrade from first to second generation antipsychotics) and so if the decision is negative, patients will not be denied treatment entirely, but rather will be offered a (perhaps only slightly) less effective treatment.

Thus, while the framework of this paper is not always appropriate, it will be in a wide range of common decision making settings.

Another possible setting where one might consider a health intervention "small" is when the intervention has a small chance of a considerable health benefit. However, how to take equity into account in situations where outcomes are risky remains a philosophically contentious issue, raising deep issues in the foundations of decision theory and welfare economics (see e.g. Fleurbaey, 2010). For the purposes of our present discussion, let us schematically distinguish an *ex ante* and *ex post* view of the healthcare system: someone who takes the *ex ante* view sees the role of the healthcare system as being to improve the quantity and distribution of chances at health; someone who takes the *ex post* view sees the role of the system as being to improve the quantity and distribution of actual realised health outcomes. Thus, in the *ex ante* view, a small chance of a considerable health benefit is indeed a small (expected) health benefit to the individual; but someone who takes the *ex post* view would refuse to accept this line of reasoning, observing that the individual will either experience the considerable health benefit or none at all, and in no event will they actually experience a small health benefit. Thus, insofar as small chances of considerable benefits are considered "small" under the *ex ante* view, and not under the *ex post* view, the *ex ante* view is more compatible with our framework presented here. To see how this might be relevant, consider that many public health interventions will provide small increases in the chances of considerable health benefits (e.g. decreasing the salt content of bread will slightly decrease the risk of stroke for everyone in the consuming population); thus taking an *ex ante* view of equity may make it easier to justify considering decisions about investments in such interventions within our framework.

To see how this framework might be used, consider a social planner who has to choose between possible healthcare investment alternatives $\mathcal{M}=\{1, \dots, i, \dots M\}$ by choosing $x = (x_1, \dots, x_i, \dots x_M)$ in some feasible set $X \subseteq [0, 1]^M$. $x_i = 0$ represents the situation where some particular alternative i is unfunded, $x_i = 1$ where it is fully funded, and any intermediate values represent the situation where it is partially funded. Associated with each investment decision x_i is a set of health improvements δ_{ij} for each individual j . Suppose that health

benefits from different alternatives combine additively so that individual j 's total health consequent on the allocation will be $h_j + \sum_{i \in \mathcal{M}} \delta_{ij} x_i$. Note that if decision variables are constrained to be binary, horizontal equity judgements may be embedded in this structure in the problem itself, in the sense that the social planner may choose to consider "Cholinesterase inhibitors for Alzheimer's disease" as a single investment alternative, or to split it into "Cholinesterase inhibitors for early to middle stage Alzheimer's" and "Cholinesterase inhibitors for late stage Alzheimer's": the latter case represents a judgement that it would be acceptable to prescribe in the early to middle stages only, and withdraw the drug for patients in the later stages.

Given the approximation (2), the social planner's problem can be written as:

$$\max_x \sum_{j \in \mathcal{N}} \alpha_j \sum_{i \in \mathcal{M}} \delta_{ij} x_i \text{ such that } x \in X \quad (3)$$

where the weights $\alpha_j = \frac{du(h_j^0)}{dh} w_j$. We stress that a higher α_j for individual j may be driven either by $\frac{du(h_j^0)}{dh}$ (reflecting poor health) or by w_j (reflecting individual characteristics). This is an important analytic distinction. For example, although government policy in England has recently prioritised health gains for people of lower socio-economic status, our reading of the relevant policy documents is not that socio-economic status itself is taken to have moral significance. Rather it is that people in the lower socio-economic classes have poorer expected health, and it is this which is morally objectionable.

In this paper we will consider two distinct versions of (3): in the continuous version, each entry, x_i takes a value in the closed interval $[0, 1]$; in the discrete version, each entry x_i takes a value in the discrete set $\{0, 1\}$. The choice between discrete or continuous variables may depend, for example, on the underlying healthcare technology (pharmaceutical provision for a particular population can be implemented at any level between full and zero, but purchasing a fraction of a dialysis centre may not be possible). Of course mixed versions with both continuous and discrete versions are possible, but to discuss this case specifically would raise no new interesting issues. The relative appropriateness of modelling healthcare investment decisions as continuous versus discrete variables has been a source of contention in the literature (Birch and Gafni, 1992; Johannesson and Weinstein, 1993). We do not take a view on the appropriateness of one relative to the other in general, but consider that each is appropriate in at least some situations, and this is sufficient justification for discussing both. In any case, much of the time our discussion applies to both versions and so we will be purposely non-specific about the domain of the variables.

It is quite simple to reframe HR-SWF (3) in terms of aggregates rather than individuals. Let the index set of subpopulations be $\mathcal{P} = \{1, \dots, k, \dots, P\}$ and write the index sets of members of the subpopulations as $\mathcal{N}_1 = \{1, \dots, N_1\}; \mathcal{N}_2 = \{N_1 + 1, \dots, N_2\} \dots \mathcal{N}_P = \{N_{P-1} + 1, \dots, N_P\}$ for suitable numbers $0 < N_1 < N_2 < \dots < N_P$ where $.$ Then (3) is equivalent to:

$$\max_x \sum_{k \in \mathcal{P}} \sum_{j \in \mathcal{N}_k} \alpha_j \sum_{i \in \mathcal{M}} \delta_{ij} x_i \text{ such that } x \in X \quad (4)$$

If all members j of subpopulation k have the same weight A_k then (4) can be rewritten as

$$\max_x \sum_{k \in \mathcal{P}} \sum_{j \in \mathcal{N}_k} A_k \sum_{i \in \mathcal{M}} \delta_{ij} x_i \text{ such that } x \in X \quad (5)$$

then by factoring out the A_k and interchanging summation signs, we get

$$\max_x \sum_{k \in \mathcal{P}} A_k \sum_{i \in \mathcal{M}} \sum_{j \in \mathcal{N}_k} \delta_{ij} x_i \text{ such that } x \in X \quad (6)$$

and if we rewrite $\Delta_{ik} = \sum_{j \in \mathcal{N}_k} \delta_{ij} \forall i, k$ then (6) can in turn be rewritten as

$$\max_x \sum_{k \in \mathcal{P}} A_k \sum_{i \in \mathcal{M}} \Delta_{ik} x_i \text{ such that } x \in X \quad (7)$$

As it stands this model is unlikely to be usable to support resource allocation in practical settings. For one thing, it is unlikely that a social planner who was prepared to commit to a determinate w_j . For another, any choice of $u(\cdot)$ would also be contentious. And even if a function $u(\cdot)$ could be agreed upon, equation (7) additionally requires the statement of the derivatives $\frac{du(h_j^0)}{dh}$ and hence in general knowledge of the complete distribution of population health h^0 . Hence there is likely to be uncertainty about the weights α_j and thus A_k . The purpose of this paper is to propose some ways in that this uncertainty might be handled.

3. Multicriteria optimisation

3.1. Basic concepts

MCO is a generalisation of mathematical programming in which, rather than a unique objective function, there are multiple, possibly conflicting "criteria functions". MCO falls within the general intellectual domain variously known as Multicriteria Decision Making, Multicriteria Decision Analysis and Multicriteria Decision Aid, henceforth referred to as MCDA (for review volumes see Belton and Stewart, 2002 and Figueira et al., 2005).

MCDA is itself not a new idea in health economics. Indeed, multicriteria techniques have been used in order to prioritise healthcare investments on a more or less *ad hoc* basis for many years - see reviews by Mullen (2004), Devlin and Sussex (2011), Airoldi and Morton (2011) and Morton et al. (forthcoming). Generally such techniques have taken the form of some sort of additive scoring and weighting method, either *ad hoc*, or within a formal framework such as Multiattribute Utility Theory.

A common approach to incorporating inequality aversion in such scoring models is to define an equity criterion, as for example in Wilson et al. (2006) and Airoldi et al. (2011), alongside a health benefit or effectiveness criterion, to assess performance on both criteria (possibly among others), weight criterion scores and sum up. This may not be an ideal approach in general, because equity, considered as a criterion, may plausibly not be preferentially independent of other criteria, in the sense that how much one cares about a change in equity may depend on the levels of other criteria. Such violations of preferential independence violate the core assumptions of the additive value model. To see the limitations of this approach, consider the example of Baltussen and Niessen (2006). This model has a "cost-effectiveness" criterion, a "severity of disease" criterion, and two equity criteria, based on whether the disease is a disease of the poor and whether it is a disease of the young respectively. The criterion "disease of the poor" in particular has a weighting of 40, indicating that according to their (illustrative) numbers, being designated a treatment for a "disease of the poor" is worth 40 points in their value function. However, one could plausibly argue that how many points one is prepared to give to a treatment because it is a "disease of the poor" should depend on whether a treatment is considered to be effective or not. In the limit, one presumably would not be prepared to give 40 points to a treatment for a disease of the poor that was completely ineffective and discredited (for comparison, in Baltussen and Niessen's table, this would far exceed the 5 points which the undoubtedly worthwhile "inpatient care for acute schizophrenia" option earns).

An alternative approach to handling inequality aversion through capturing equity as a criterion in an MCDA model is to consider the health gains to different population groupings themselves as the multiple criteria. In such an approach it is easy to relate the value model to formula (7). This overcomes the theoretic difficulty about criterion independence. However, as noted above, a drawback from the point of view of practical elicitation is that social planners are likely to be extremely reluctant to state tradeoffs between different population subgroups directly. Hence an approach that can deal with uncertainty about the weight parameters is required.

This leads us into our discussion of MCO. MCO problems are typically written in the following notation, where $\mathcal{P} = \{1, \dots, P\}$ is the index set of a set of criteria that a decision maker may wish to maximise. Observe that the critical difference from a standard mathematical program is that there are multiple criteria functions, separated by commas, instead of a single objective function.

$$\max_x W_1(x), \dots, W_P(x) \text{ such that } x \in X \quad (8)$$

We will write the vector $(W_1(x), \dots, W_P(x))$ as $W(x)$.

In the multicriteria setting, one way to define the concept of efficiency is as follows.

Definition 1. *A solution x' is efficient if there is a monotonically increasing functional $v(\cdot)$ such that $v(W(x)) \leq v(W(x')) \forall x \in X$.*

This is slightly different from the more familiar definition for x' to be efficient (there has to be no solution x such that $W_k(x) \geq W_k(x') \forall k \in \mathcal{P}$ and $\exists k^o: W_{k^o}(x) > W_{k^o}(x')$) but can be easily seen to be equivalent (to see that our definition implies the usual definition, observe that if our definition holds but not the familiar definition, $v(\cdot)$ cannot be increasing which is a contradiction; if the familiar definition holds then we take $v(\cdot) = W_{k^o}(\cdot)$ and so our definition must hold). The reason for using this non-standard definition will become clear in the ensuing.

For our purposes, however, this definition seems excessively general. We are not interested in all possible functionals, but only those functionals that could be reasonably described as inequality averse, and Definition 1 includes functionals that model both inequality averse and inequality seeking preferences.

A refinement of the concept of Definition 1 is to specify efficiency relative to the unit simplex $\mathcal{A}_o = \{(A_1, \dots, A_P) \in \mathbb{R}_{++}^P : \sum_{k \in \mathcal{P}} A_k = 1\}$ (the restriction that weights sum to 1 is an arbitrary normalisation).

Definition 2. *A solution x' is convex efficient if there is a set of weights $\{A_1, \dots, A_P\} \in \mathcal{A}_o$ such that $\sum_{k \in \mathcal{P}} A_k W_k(x) \leq \sum_{k \in \mathcal{P}} A_k W_k(x') \forall x \in X$.*

Solutions that are efficient but not convex efficient are called "unsupported" efficient solutions. It is well-known that in cases where the feasible set is convex the set of efficient and convex efficient solutions coincides (Ehrgott, 2005). Hence, if the feasible set is convex, there are no unsupported efficient solutions.

A further refinement of the concept of efficiency is to specify efficiency relative to the linear additive functional with weights in some given set $\mathcal{A} \subseteq \mathcal{A}_o$.

Definition 3. *A solution x' is \mathcal{A} -efficient if there is a set of weights $\{A_1, \dots, A_P\} \in \mathcal{A}$ such that $\sum_{k \in \mathcal{P}} A_k W_k(x) \leq \sum_{k \in \mathcal{P}} A_k W_k(x') \forall x \in X$.*

Thus, the critical difference between Definition 2 and Definition 3 is that the former uses the whole of the weight simplex \mathcal{A}_o and the latter uses a specific subset of the weight simplex, \mathcal{A} . Observe that since all \mathcal{A} -efficient solutions are convex efficient, no \mathcal{A} -efficient solutions are unsupported.

This refinement is particularly relevant in our case, as some groups may be clearly disadvantaged relative to others, and so may attract higher weights. Hence, what may be useful to a social planner interested in solving (7) but unsure of the values of the A_k s is the \mathcal{A} -efficient solutions to a problem of the form:

$$\max_x \sum_{i \in \mathcal{M}} \Delta_{i1} x_i, \dots, \sum_{i \in \mathcal{M}} \Delta_{iP} x_i \text{ such that } x \in X \quad (9)$$

where \mathcal{A} is defined by a set of constraints on weights (for example, perhaps the weight associated with population 1 may be less than or equal to the three quarters of the weight associated with population 2). As it happens, algorithms

for finding \mathcal{A} -efficient solutions to MCO problems have been an active area for research in recent years (Liesiö et al., 2007; Argyris et al., 2011).

3.2. Non-supported efficient solutions

An example of efficient but non-convex efficient solutions can be constructed in settings where the dimensions are gains to different population groups. In the example of Figure 1, the social planner aims to maximise health gains to population groups 1 and 2 and there are three discrete solutions a , b and c , all of which are efficient, but only a and b are convex efficient: any linear isopreference line will pick out a or b as superior to c .

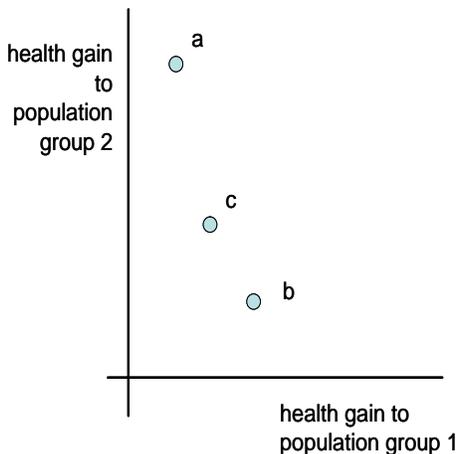


Figure 1. Three efficient solutions, two of them convex efficient

There *is* a rationale for considering point c as "better than" a or b , in the sense that it is a "good compromise". Indeed, if the population groups are sociologically meaningful entities, the members of which perceive themselves as having common interests and the ability to undertake common actions (rather than purely statistical constructs, such as socio-economic groups), then one could view c as the likely outcome of some sort of bargaining process between these entities. However, this seems to lead to a view of healthcare prioritisation as a (perhaps implicit) negotiation between different sections of society (the reader is referred to Hauck et al. (2004) for a fuller development of this perspective). This choice is, however, incompatible with the analysis developed in this paper, as such a solution would not be an \mathcal{A} -efficient solution to (9).

In the light of this analysis, it is interesting to consider the practice of imposing equity constraints directly, as recommended by Stinnett and Paltiel (1996) and others in the mono-criterion environment. To fix ideas consider again the case where there are two population groups. In that case, (9) specialises to:

$$\max_x \sum_{i \in \mathcal{M}} \Delta_{i1} x_i, \sum_{i \in \mathcal{M}} \Delta_{i2} x_i \text{ such that } x \in X \quad (10)$$

In the case where population group 1 was worse off than population group 2, in our approach this could be captured in a weight constraint set $\mathcal{A} = \{(A_1, A_2) : (A_1, A_2) \in \mathbb{R}_{++}^2, A_1 + A_2 = 1 \text{ and } A_1 \geq A_2\}$. In contrast, in the mono-criterion environment, when equity constraints are used, this concern about equity would be captured by adding an additional constraint on the decision variables, and the relevant math program would be:

$$\max_x \sum_{i \in \mathcal{M}} \Delta_{i1} x_i + \sum_{i \in \mathcal{M}} \Delta_{i2} x_i \text{ such that } x \in X \text{ and } \sum_{i \in \mathcal{M}} \Delta_{i1} x_i \geq B \quad (11)$$

It is quite easy to show that, in the discrete version of this problem, there can be values of B for which (11) has solutions which are not \mathcal{A} -efficient with respect to (10). Figure 2 demonstrates this using the example of Figure 1. As noted above, c is not convex efficient, and hence cannot be \mathcal{A} -efficient. However, if B is set as shown on the diagram (and so a is "cut off") and the (linear) isoquant of the objective function is the diagonal line through c then c will be chosen as an optimal solution to (11) over b . Hence, the policy recommendations that fall out from the use of equity constraints cannot be justified within our framework. Choices about how to model equity (e.g. about whether to model equity in the constraints or in the objectives of a mathematical program) seem to be made on quite casual grounds: it is important in our view that analysts should understand that different models can lead, not only to different solutions, but to different sorts of solutions.

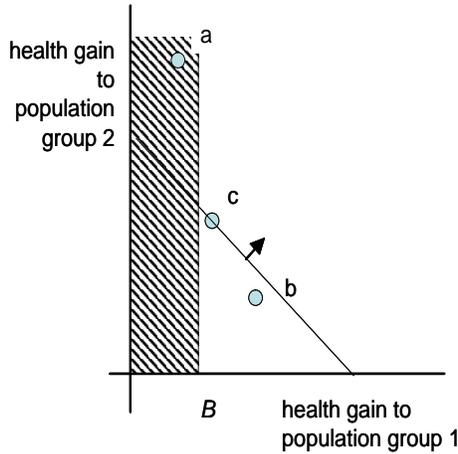


Figure 2. c is optimal but not \mathcal{A} -efficient

3.3. Relation to classical view of economic inequality

There is a classical literature on the comparison of inequal distribution of economic goods (such as income and health), motivated by the need to make statements such as "income is distributed more unequally in the United States than in the United Kingdom", and the aim of the current section is to compare our framework with the ideas in that literature. The classical line of research,

beginning with Atkinson (1970) and Kolm (1969), and given its definitive statement in the first edition of Sen (1997), seeks to understand the comparison of different distributions of income through an analogy with the comparison of gambles (i.e. distributions of probability mass). Intuitively, a risk averse decision maker will like a monetary gamble more (less) if the probability mass is concentrated (spread out), so that the expected value of the gamble remains the same but there are lower (higher) chances of both better and worse outcomes; in a similar way, an inequality averse social planner will like an income distribution more (less) if the same amount of income is concentrated (spread out) by transferring income from richer to poorer (poorer to richer) individuals. This sort of observation gives a way to make inequality comparisons between different distributions of income.

The clearest statement of the seminal result in this literature is the paper of Dasgupta, Sen and Starrett (1973). A slightly simplified version of what these authors show is the following. For some real-valued total income I and some natural number n of people, define $Z(I)$ as the set of possible ordered distributions I , i.e. vectors $z = (z_1, \dots, z_n)$ with $z_1 \leq \dots \leq z_n$ and $\sum_{i=1}^n z_i = I$. For z' and $z'' \in Z(I)$, the following conditions are equivalent: (i) for all strictly Schur-concave social welfare functions $v(\cdot)$, $v(z') > v(z'')$; (ii) the Lorenz curve (showing cumulated income) associated with z' is everywhere (nonstrictly) above the Lorenz curve associated with z'' ; (iii) z' can be obtained from z'' by a series of ("Pigou-Dalton") transfers of income from rich to poor. These three conditions provide different perspectives on the statement " z' is a more equal distribution of income than z'' ": (i) relates the concept of inequality to an abstract mathematical concept, that of the Schur-concavity of a function; (ii) provides a computational way to test for two distributions whether one is definitely more unequal than another; and (iii) gives a constructive perspective, showing how we might move from a less to a more equal distribution.

Extensions and variations of this approach have been widely studied in both general economics and in health economics. For example in the economic mainstream, Shorrocks (1983) shows how to compare distributions with different sum total incomes using Generalised Lorenz curves. In health economics, Wagstaff (1991, 2002) discusses the use of "concentration curves" in the place of Lorenz curves to determine which of two distributions are more unequal in the sense that richer people are more healthy and poorer people are less healthy; Allison and Foster (2006) and Sonne-Schmidt et al (2013) discuss how to compare distributions where only ordinal data relevant to well-being is available, by defining improving (degrading) transfers of population between ordinal categories.

To highlight the relationship between this sort of approach and our approach in the current paper, it is instructive to define an efficiency concept along the lines of condition (i) of Dasgupta, Sen and Starrett. This definition is clearly analogous to our Definitions 1-3, and comparing Definition 3 and Definition 4 may be a useful way of understanding the distinctive features of our approach in relation to that of the classical economic inequality literature.

Definition 4. A distribution z' is DSS-efficient if there is a strictly Schur-concave functional $v(\cdot)$ such that $v(W(z)) \leq v(W(z')) \forall z \in Z(I)$.

We highlight three differences between Definitions 1-3 and Definition 4.

- Firstly, Definition 4 contains no mention of the increasingness of the function $v(\cdot)$. One way to look at this is to observe that as the result is restricted to distributions of some determined amount I of total income, there will not be two distributions z' and $z'' \in Z(I)$: $z' \geq z''$. Therefore, in definition Definition 4 the issue of increasingness or otherwise of $v(\cdot)$ does not arise. As noted, Shorrocks (1983) and other authors have explored the situation where the assumption of constant sum total income is relaxed.
- Secondly, the framework of Dasgupta, Sen and Starrett embeds a symmetry assumption in the sense that all Schur-concave functions will assign the same value to all permutations. In our framework, we make no such symmetry assumption: indeed in our original statement of the HR-SWF in equation (1), we explicitly allow that the health of members of the population may be weighted differently (e.g. because their poor health is self-inflicted). If all weights w_j are the same, then α_j effectively collapses to $\frac{du(h_j^0)}{dh}$ (possibly multiplied by an irrelevant constant), and since $u(\cdot)$ is concave by assumption.
- Thirdly (unlike for Dasgupta et al.), we want to move beyond concepts of the type of Definition 1 and Definition 4, to concepts of the type of Definition 3, that allow us explore how the implications of arbitrary weight sets. This is likely to be useful since while decision makers may be unable to assign precise tradeoff weights to health gains for different population groups, they may be quite able to reject both grossly inequality-sensitive and inequality-insensitive weights, and hence make statements such as "a health gain to the most deprived population subgroup is worth between 120 and 140% of the same health gain to the least deprived population subgroup".

4. Example

4.1. The resource allocation problem

In this section, we present a numerical example to illustrate the concepts developed above. The example is based on a model of treatment for depression, described in Morton et al. (2008). Two features of depression are salient. One is that is that the temporal characteristics of depression are very variable: some sufferers may experience a single episode of depression and then recover fully, while others may experience recurrent episodes, while still others may experience an extended, unremitting chronic depressive episode. One of the challenges in treating depression cost-effectively is therefore to find ways to target resources

to sufferers whose illness will not spontaneously remit without treatment, hence the popularity of "stepped care" models (NICE, 2004). Another relevant feature of depression is its relatively high population prevalence - e.g. 2.6% in the UK according to the Psychological Morbidity Survey of 2000 according to Singleton et al (2001) - and the presence of significant unmet need for care (Andrews and Henderson, 2000).

Depression has three severity levels, and these three severity levels are crossed with the transitory/ recurrent/ chronic distinction suggested above, to give nine population groups in total. For an inequality averse social planner, treatments that benefit those in the most unfortunate group should be prioritised. However, while it is clear that, for example moderate recurrent depression is worse than moderate transitory depression, it is not clear how it stands relative to severe transitory depression. Hence the population groups cannot be ordered, but they can be quasi-ordered as shown in Figure 3 (population groups will henceforth be referred to using the 3-letter codes on the figure). The intended interpretation of Figure 3 is that *severe chronic* is worse than *severe recurrent* and *moderate chronic* depression and so on. We define \mathcal{A} by requiring that the weight attached to alternatives that benefit worse off patients is constrained to be higher than the weight for less badly-off patients, requiring weights across the nine patient groups to sum to 1, and requiring that the weight assigned to alternative projects benefitting sufferers of mild transitory depression is at least 1/12 (bearing in mind that weights across all nine categories must sum to unity). The use of 1/12 as a lower bound for the weight of the lowest priority group is purely illustrative, and other numbers could be chosen.

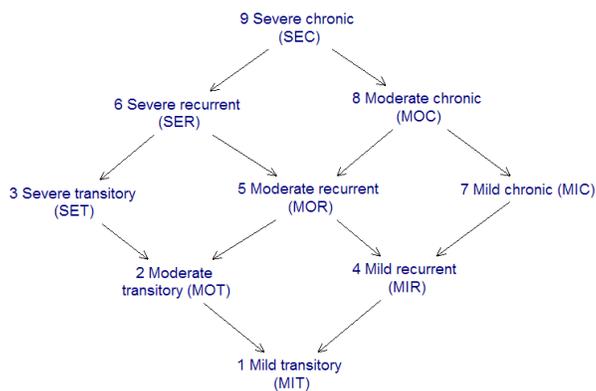


Figure 3. Ordering of depression subgroups

We consider three sorts of mutually exclusive alternatives for each population group: improving coverage, improving the treatment mix as per NICE guideline at current treatment levels, and improving both coverage and treatment mix (with the exception of MIT, for whom we only model extending coverage;

appropriate treatment for this group consists primarily of watchful waiting and so there is little to be done with improving treatment mix). The problem is thus one for which there are multiple clusters of mutually exclusive interventions (Johannesson and Weinstein, 1993). Where decisions variables are discrete this is known in the operations research literature as the multiple choice knapsack problem (Kellerer et al., 2004). Costs and benefits of these alternatives as shown in Table 4. Benefit estimates draw on modelling described in Morton et al. (2008); costs draw on Secta (2004). We will suppose that costs combine additively and there is a fixed budget available for expenditure, i.e. that $\sum_{i \in \mathcal{M}} c_i x_i \leq B$ for some budget B .

	Benefits (QALY 000s)			Costs (£ms)		
	Mix	Cov	Both	Mix	Cov	Both
1. Mild transitory (MIT)	-	1.2	-	-	13.7	-
2. Moderate transitory (MOT)	0.1	4	4.2	2.6	13.3	17.7
3. Severe transitory (SET)	0	3	3.1	2	5.1	8.3
4. Mild recurrent (MIR)	0	0.4	0.5	10.1	9.5	26.3
5. Moderate recurrent (MOR)	0.3	1.7	2.1	15	14.1	39.2
6. Severe recurrent (SER)	0.6	3.5	4.5	9.9	10.6	27.1
7. Mild chronic (MIC)	0.1	0.3	0.5	4.1	2.8	9.8
8. Moderate chronic (MOC)	0.3	0.9	1.3	5	3.4	11.7
9. Severe chronic (SEC)	0.2	0.8	1.3	2.4	1.6	5.6

Table 4. Costs and benefits for investment alternatives for depression

Hence, we have all the components of a MCO problem of the form (9). We will consider now consider how to characterise the \mathcal{A} -efficient solutions in both continuous and discrete versions of the problem.

4.2. Analysis of the continuous version

From a formal point of view, we are interested in finding \mathcal{A} -efficient solutions to the MCO problem, which is a specialisation of problem (9).

$$\begin{aligned}
 & \max_x \sum_{i \in \mathcal{M}_1} \Delta_i x_i, \dots, \sum_{i \in \mathcal{M}_P} \Delta_i x_i \\
 & \text{such that} \\
 & \sum_{i \in \mathcal{M}} c_i x_i \leq B \quad \text{(i)} \\
 & \sum_{i \in \mathcal{M}_k} x_i = 1 \quad \forall k \in \mathcal{P} \quad \text{(ii)} \\
 & x \in [0, 1]^M \quad \text{(iii)}
 \end{aligned} \tag{12}$$

In this formulation, $\mathcal{P} = \{1, \dots, P\}$ with $P = 9$ and $\mathcal{M} = \{1, \dots, M\}$ with $M = 25$ are the index sets for the population groups and alternatives respectively. \mathcal{M}_1 through \mathcal{M}_P is a partitioning of \mathcal{M} into alternatives consumed by each of the population groups. Constraint (i) is the budget constraint; constraints (ii) ensure that only one alternative is chosen for each population group; constraint (iii) ensures the decision variables lie in $[0, 1]$.

For each weight vector α a family of solutions to (12) for all values of B can be generated by an algorithm described in Johannesson and Weinstein (1993). This algorithm is proved optimal in Section 1 of Sinha and Zoltners (1979), a fact which seems not to be known in the health economics literature (Johannesson and Weinstein cite a 1991 unpublished PhD thesis in Swedish as their authority). There are three stages to the algorithm: (i) the elimination within cluster of dominated alternatives, (ii) ordering of the remaining alternatives within cluster by cost and computation of the incremental cost effectiveness ratios, and finally (iii) ranking of the resulting incremental increases in provision. Since the clusters are population groups, and weights are specific to population groups, the dominance relation does not depend on the choice of α and we are able to carry out steps (i) and (ii) in the usual manner. (In fact, elimination of dominated alternatives leads to the elimination of all the "Mix" alternatives, leaving 17 alternatives, two for each population group except MIT.)

Carrying out step (iii) without knowledge of which α is to be applied is problematic, however. Hence, we turn the problem around, asking, for any pair of increments i' and i'' is there a choice of α which leads to i' being less cost-effective than i'' ? This can be done by solving the following linear program:

$$\min_{\alpha} \alpha_{k(i')} \Delta_{i'k} / c_{i'} - \alpha_{k(i'')} \Delta_{i''k} / c_{i''} \text{ such that } \alpha \in \mathcal{A} \quad (13)$$

In the event that the value of (13) is greater than 0, then it is impossible for increment i' to be ranked lower than increment i'' by any choice of α . Performing pairwise tests in this manner generates a quasi-ordering of alternatives: this can be interpreted as the intersection quasi-ordering for all possible cost-effectiveness orderings for all $\alpha \in \mathcal{A}$. Solving such linear programs is straightforward and was done in MSExcel using the solver add-in. The resulting quasi-ordering for the depression example is shown as Figure 4.

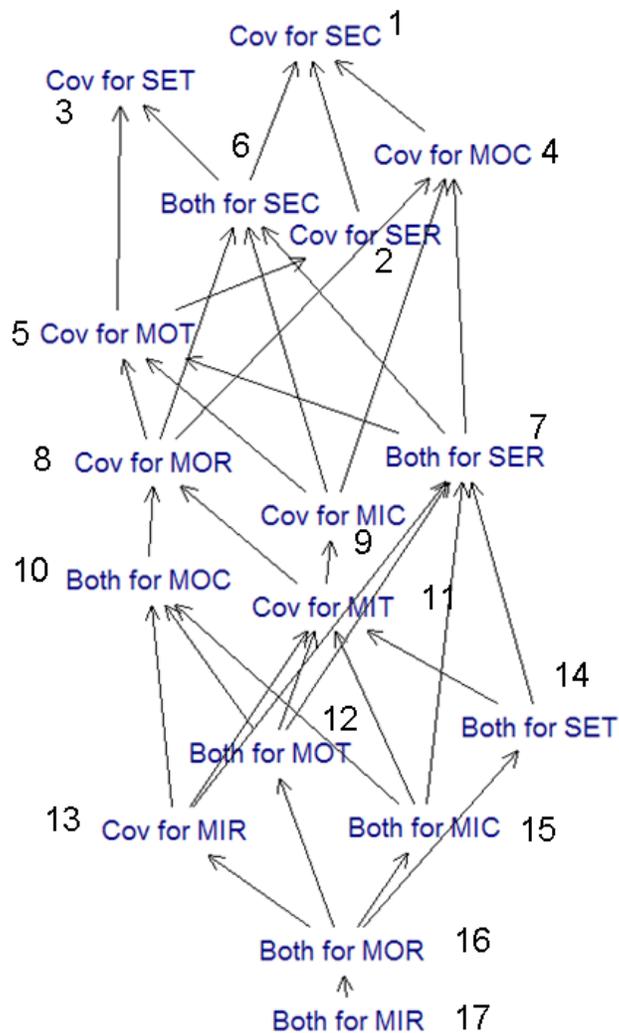


Figure 4. Quasi-ordering of alternatives for treating depression

The interpretation of the figure is that an alternative i' is superordinate to another alternative i'' in the order (it is possible to trace a sequence of tail-head arrows going from i'' to i') if and only if it has a higher benefit-cost score for all feasible weights in the set \mathcal{A} . Thus, all \mathcal{A} -efficient solutions to the continuous MCO resource allocation problem (12) can be generated by starting at one of the heads of the graph (the nodes with no outgoing arrows, either *improve coverage for SET* or *improve coverage for SEC*) and proceeding down the graph selecting increments in such a way that no increment is selected before all the increments superordinate to it in the graph until funds are exhausted. In summary, insofar as alternatives are more superordinate or "higher up" on this figure, they are more robust to variations in the weights attached to the

different population groups because they will always be prioritised earlier than alternatives less superordinate or "lower down" in the ordering.

To make this clear, the numbering of the alternatives in Figure 4 shows the order in which increments will be selected when weights are 1/6 for population groups SEC, SER and MOC, and 1/12 for all other population groups. The reader can verify that no increment is selected before all the increments superordinate to it are selected (e.g. "Cov for MOC" is not selected until "Cov for SER", "Cov for SEC" and "Cov for SET" are selected).

4.3. Analysis of the discrete version

We now turn our attention to the \mathcal{A} -efficient solutions of the problem of the form:

$$\begin{aligned}
 & \max_x \sum_{i \in \mathcal{M}_1} \Delta_i x_i, \dots, \sum_{i \in \mathcal{M}_P} \Delta_i x_i \\
 & \text{such that} \\
 & \sum_{i \in \mathcal{M}} c_i x_i \leq B \quad \text{(i)} \\
 & \sum_{i \in \mathcal{M}_k} x_i = 1 \quad \forall k \in \mathcal{P} \quad \text{(ii)} \\
 & x \in \{0, 1\}^M \quad \text{(iii')}
 \end{aligned} \tag{14}$$

This is identical to problem (12) except that constraint (iii') ensures that the decision variables cannot take fractional values.

As is well-known, the mono-criterion version of the problem (14), even without the multiple choice constraints (ii), is NP-hard. Although customised algorithms to solve this problem do exist, hand solution is generally impractical for all but the simplest cases. Moreover, as the budget increases, items may both enter and leave the optimal solution, hence there is no analogous ordering to the cost-effectiveness ordering of the previous section.

In general, as in the mono-criterion environment, solving combinatorial MCO problems is more difficult than solving linear MCO problems, requiring both greater algorithmic ingenuity and computer processing time. Indeed, while algorithms for multicriteria linear programs have been well-known for some time (Steuer, 1986), algorithms for multicriteria integer and combinatorial problems are very much a live research field (Ehrgott and Gandibleux, 2000; Alves and Clímaco, 2007). To identify the \mathcal{A} -efficient solutions for problem (14) we use a computer code which implements the algorithm described in Argyris et al. (2011). The idea of the algorithm is form a single combinatorial optimisation problem with both choice of alternative x and weights A_k variable, and solve multiple times, progressively cutting off solutions after they have been identified.

We have run this algorithm for a budget constraint of £100m and show the results of the analysis in Figure 5. For each population group, the chart shows the proportion of \mathcal{A} -efficient solutions that includes each of the "mix", "coverage" or "both" alternatives (bearing in mind that only one alternative can be selected for each population group). The results are similar to the results

of the analysis of the continuous version of the problem (for example, note that in no \mathcal{A} -efficient solution is the dominated "mix" alternative selected; for the population groups MIR and MOR for whom upgrading from the "coverage" to the "both" alternative is highly cost-ineffective, that upgrade is not included in any \mathcal{A} -efficient solution). This display is a variant of the core index display of Liesiö et al (2007), although that display is designed for a context without the multiple choice constraints (ii) of (14).

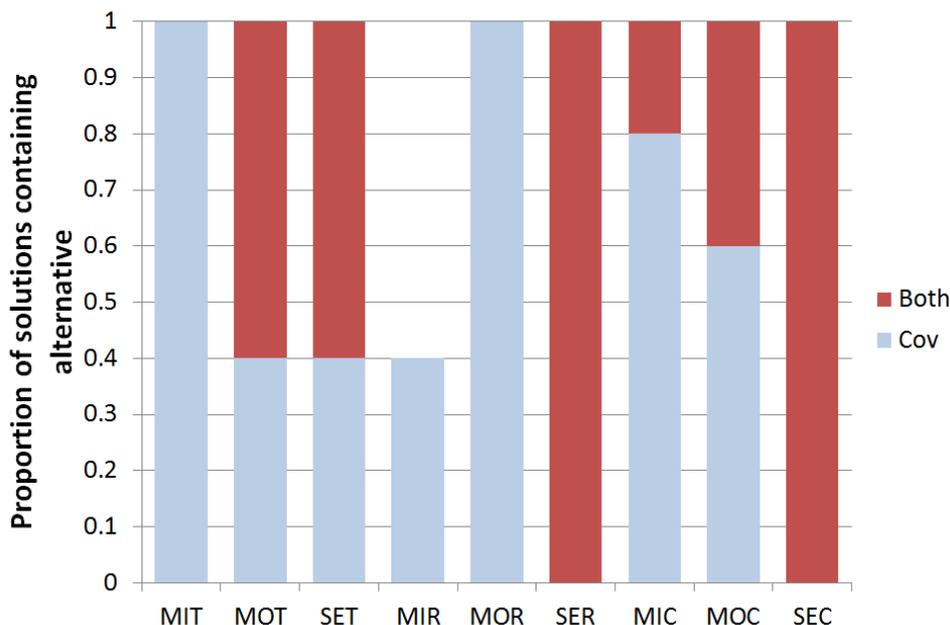


Figure 5. Proportion of time a solution is included in an \mathcal{A} -efficient solution

It can be seen that even with this relatively limited information about the weights of health gains to different populations, it is possible to arrive at quite crisp conclusions about what to do: one should implement the "both" alternative for SER and SEC and the "coverage" alternative for MIT and MOR. Thus, out of nine population groups for whom one has to make a decision, the model gives recommendations about what to do for four of them. In an application context, this could be quite useful to a decision maker who would then be able to focus attention on the value tradeoffs that bear directly on the undecided population groups.

We note that a similar analysis can be done in the continuous case, as described in the previous section. One way to do this is randomly simulate weights from \mathcal{A} . This can be done by using standard methods to simulate weights uniformly from the unit simplex (using algorithm 2 of Onn and Weisman, 2011), then rejecting (Ross, 2002) weight vectors that do not fulfill the stipulated weight constraints. Then one can use each set of simulated weights to

generate a priority ordering, and allocate resources by proceeding down the priority ordering until the budget is exhausted. Finally, the proportion of simulation iterations for which the coverage and both options are implemented (as the mix option by itself is dominated in all cases) can be calculated for each population subgroup. For 500 iterations, the resulting chart is shown in Figure 6.

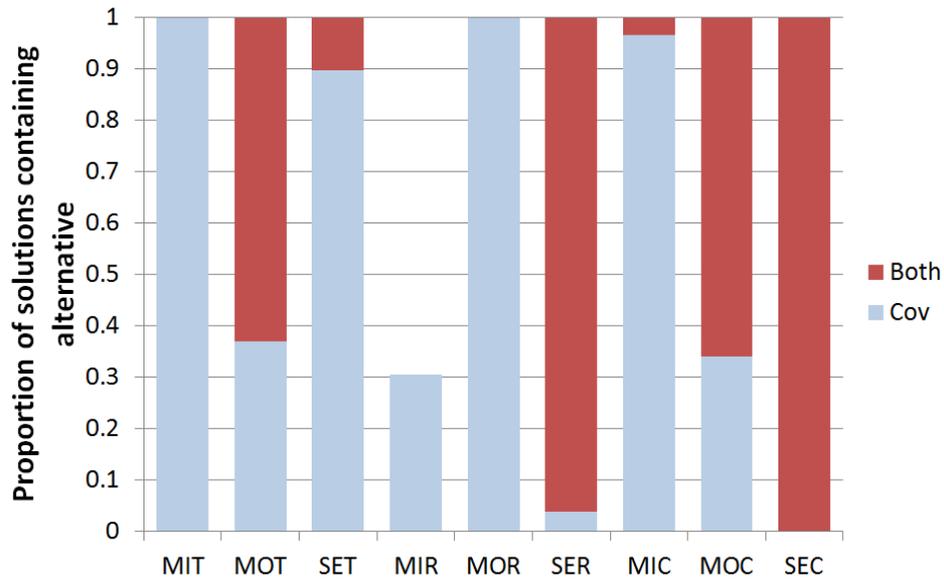


Figure 6. Proportion of time an alternative is included in a simulated optimal solution

It can be seen that while the specific numbers in Figures 5 and 6 are different, the qualitative messages are similar. For example, implementing coverage for MIT and MOR and both for SEC is definitely a good idea according to both displays, and implementing both for SER is certainly worthwhile under Figure 5 and almost certainly worthwhile in Figure 6. The only population subgroup that might be left without any change either in intervention mix or coverage is MIR in both charts. Hence these figures underscore that what is critically important in analysis is the problem frame, the data, and the constraints that are placed on social values: the fine technical details of the formulation of the model and solution procedure have rather minimal bearing on the resulting decision recommendations.

5. Conclusion

In this paper we have outlined a formal model of the healthcare resource allocation problem, drawing on the welfare economic theory of health. We have

shown how the key uncertainties about values which make parameterising such models problematic can be easily handled within a multicriteria formulation. We have moreover demonstrated how such formulations might be applied in practice using an example based on the allocation of funds for the treatment of various forms of depression.

Some years ago, Smith (1995) argued that "far from being a strength, the stark clarity of mathematical models may make them unacceptable to the accountable person" (p 154), forcing the decision maker to make unacceptably explicit value judgements. Understanding relevant tradeoffs is a hallmark of rational decision and we do not wish to suggest that responsible analysts should hide from decision makers the consequences of their choices. Nevertheless, an important insight of Smith's paper is that an excessive focus on specifying the maximand can distract from making decisions and engender needless conflict - to take a trivially simple but nevertheless instructive case, a decision maker faced with a choice between giving an equal benefit (such as a QALY) to person A or person B only needs to know which of the two candidates is more deserving, not quantitatively how much more deserving they are.

The advantage of the formulations presented in this paper is that they can be used, at best to help decision makers make rational, efficient choices in a situation where provision of precise parameters expressing key value tradeoffs is difficult; and at worst, to focus attention on the key, decision-relevant, value tradeoffs at stake. Moreover, our approach relates more naturally to the underpinning welfare economic theory of healthcare resource allocation than approaches based on specifying equity constraints requiring particular levels or shares of health or resources for particular population subgroups, or approaches based on defining an equity criterion amongst other criteria.

Will the set of \mathcal{A} -efficient solutions be unmanageably large and diverse? Perhaps, but under plausible conditions, this seems unlikely. Value differences that seem unbridgeable when the debate is framed as one of principle (should the healthcare system treat smokers and non-smokers differently?) may seem much more manageable when presented as questions about parameter choice (how much priority do we want to give nonsmokers over smokers in access to healthcare?). Moreover, much healthcare decision making takes place in a deliberative setting, where the key players genuinely strive to understand opposing views. In such an environment, differences in values, and the parameter intervals representing these differences in values, are likely to shrink rather than grow as discussion progresses. As parameter intervals become narrower, the number of \mathcal{A} -efficient solutions will become smaller and these solutions will become more similar.

Overall, we think the multicriteria methods that we propose in this paper have great potential in healthcare resource allocation. Healthcare is an environment where values are inevitably contested, as it concerns the distribution of goods of an unusually fundamental nature. The present paper has focussed on preferences over distribution, but there are still deeper ethical issues which remain to be explored: for example, should the maximand be a function of health at all, or capability, well-being or monetised economic activity? Multicriteria

formulations similar to those which we present in this paper have a great deal to offer health economists who wish to be able to make robust recommendations when all these perspectives point to the same decisions, and to illuminate the key tradeoffs when values do indeed conflict.

Acknowledgement 5. *This paper was informed and stimulated by the development of the STAR approach funded by the Health Foundation. The core STAR team (in addition to the author) consisted of Dr Mara Airoidi, Professor Gwyn Bevan and Dr Jenifer Smith. Additionally, the example of clinical depression developed in Section 4 was originally developed with support from the Health Foundation under the QQUIP programme. The author is especially grateful to Mara Airoidi and Özlem Karsu for many discussions, to Dr Nikos Argyris for sharing the computer implementation of the code described in Section 4.3, and to two reviewers for insightful and constructive comments.*

6. References

Airoidi M, Morton A, Smith J, Bevan G. Healthcare prioritisation at the local level: a socio-technical approach. SYMPOSE working paper no 7. Department of Management, London School of Economics and Political Science: London; 2011.

Airoidi M, Morton A. 2011. Portfolio Decision Analysis for population health. In: Salo A, Keisler J, Morton A (Eds), Portfolio Decision Analysis: improved methods for resource allocation. Springer: New York; 2011.

Allison RA, Foster JE. Measuring health inequality using qualitative data. *J Health Econ* 2004;23: 505-524.

Alves MJ, Clímaco J. A review of interactive methods for multiobjective integer and mixed integer programming. *European Journal of Operational Research* 2007;180: 99-115.

Anand P. The integration of claims to health-care: a programming approach. *Journal of Health Economics* 2003;22: 731-745.

Andrews G, Henderson S. 2000. Unmet need in psychiatry. (Eds). CUP: Cambridge; 2000.

Argyris N, Figueira JR, Morton A. Identifying preferred solutions to Multi-Objective Binary Optimisation problems, with an application to the Multi-Objective Knapsack Problem. *Journal of Global Optimization* 2011;49: 213-235.

Atkinson AB. On the measurement of inequality. *Journal of Economic Theory* 1970;2: 244-263.

Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost effectiveness and resource allocation* 2006;4.

Belton V, Stewart TJ. Multiple Criteria Decision Analysis: an integrated approach. Kluwer: Boston; 2002.

Bevan G, Airoldi M, Morton A, Oliveira M, Smith J. Estimating health and productivity gains in England from selected interventions. Health Foundation: London; 2007.

Birch S, Gafni A. Cost-Effectiveness Utility Analyses - Do Current Decision Rules Lead Us to Where We Want to Be. *Journal of Health Economics* 1992;11; 279-296.

Birch S, Gafni A. Changing the Problem to Fit the Solution - Johannesson and Weinstein (Mis) Application of Economics to Real-World Problems. *Journal of Health Economics* 1993;12; 469-476.

Cleary S, Mooney G, McIntyre D. Equity and Efficiency in HIV-Treatment in South Africa: The Contribution of Mathematical Programming to Priority Setting. *Health Economics* 2010;19; 1166-1180.

Culyer AJ. The normative economics of healthcare finance and provision. *Oxford Review of Economic Policy* 1989;5; 34-55.

Dantzig G. Linear programming and extensions. Princeton: Princeton University Press; 1998.

Dasgupta P, Sen A, Starrett D. Notes on Measurement of Inequality. *Journal of Economic Theory* 1973;6; 180-187.

Devlin NJ, Sussex J. 2011. Incorporating Multiple Criteria in HTA: Methods and Processes. (Eds). Office of Health Economics: London; 2011.

Dolan P, Edlin R, Tsuchiya A. The relative value of health gains to different beneficiaries. Final Report RM03/JH11. Report to the National Collaborating Centre for Research Methodology. <http://www.haps.bham.ac.uk/publichealth/methodology/docs/publications> Accessed 22/02/10.; 2008.

Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. *Health Economics* 2005;14; 197 - 208.

Drummond KF, Sculpher MJ, Torrance GW, O' Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. OUP: Oxford; 2005.

Earnshaw SR, Dennett SL. Integer/linear mathematical programming models - A tool for allocating healthcare resources. *Pharmacoeconomics* 2003;21; 839-851.

Ehrgott M. *Multicriteria optimisation*. Springer: Berlin; 2005.

Ehrgott M, Gandibleux X. A survey and annotated bibliography of multi-objective combinatorial optimization. *OR Spektrum* 2000;22; 425-460.

Epstein DM, Chalabi Z, Claxton K, Sculpher M. Efficiency, equity, and budgetary policies: Informing decisions using mathematical programming. *Medical Decision Making* 2007;27; 128-137.

Epstein D, Jimenez-Rubio D, Smith P, Suhrcke M. Social determinants of health: an economic perspective. *Health Economics* 2009;18; 495-502.

Figueira JR, Greco S, Ehrgott M. 2005. *Multiple Criteria Decision Analysis: State of the Art Surveys*. (Eds). Springer: New York; 2005.

Fleurbaey M. Assessing Risky Social Situations. *Journal of Political Economy* 2010;118; 649-680.

- Hauck K, Smith PC, Goddard M. The economics of priority setting for healthcare: a literature review. World Bank: Washington, DC; 2004.
- Johannesson M, Weinstein MC. On the Decision Rules of Cost-Effectiveness Analysis. *Journal of Health Economics* 1993;12; 459-467.
- Keeney RL, Raiffa H. Decisions with multiple objectives: preferences and value tradeoffs. Wiley: Chichester; 1976.
- Kellerer H, Pferschy U, Pisinger D. Knapsack Problems. Springer-Verlag: Berlin; 2004.
- Kolm SC. 1969. The optimal production of social justice. In: Guitton H, Margolis J (Eds), *Public Economics*. St Martin: New York; 1969.
- Liesjö J, Mild P, Salo A. Preference programming for robust portfolio modeling and project selection. *European Journal of Operational Research* 2007;181; 1488-1505.
- Morton A, Airoidi M. Incorporating health inequalities considerations in PCT priority setting. Operational Research Group, Department of Management. London School of Economics and Political Science; 2009.
- Morton A, Bevan G, Airoidi M, Oliveira M, Smith J. Estimating the health gains and cost impact of treatment for depression in England. Health Foundation: London; 2008.
- Mullen PM. Quantifying priorities in healthcare: transparency or illusion. *Health Services Management Research* 2004;17; 47-58.
- NICE. Depression: management of depression in primary and secondary care. NICE: London; 2004.
- Onn S, Weissman I. Generating uniform random vectors over a simplex with implications to the volume of a certain polytope and to multivariate extremes. *Ann Oper Res* 2011;189; 331-342.
- Østerdal LP. Axioms for healthcare resource allocation. *Journal of Health Economics* 2005;24; 679-702.
- Payne G, Laporte A, Deber R, Coyte PC. Counting backward to healthcare's future: using time-to-death modeling to identify changes in end-of-life morbidity and the impact of aging on health expenditures. *The Millbank Quarterly* 2007;85; 213-257.
- Ross S. Simulation. Academic Press: San Diego; 2002.
- Sen A. On economic inequality: enlarged editions. Clarendon Press: Oxford; 1997.
- Shorrocks AF. Ranking Income Distributions. *Economica* 1983;50; 3-17.
- Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer H. Psychiatric morbidity among adults living in private households. HMSO: London; 2001.
- Sinha P, Zoltners AA. Multiple-Choice Knapsack Problem. *Operations Research* 1979;27; 503-533.
- Smith P. Large scale models and large scale thinking: the case of the health services. *Omega-International Journal of Management Science* 1995;23; 145-157.
- Sonne-Schmidt C, Tarp F, Østerdal LP. Ordinal multidimensional inequality. Unpublished working paper, University of Southern Denmark, Odense.
- Steuer RE. Multiple Criteria Optimization: Theory, Computation and Application. Wiley: New York; 1986.

Stinnett AA, Paltiel AD. Mathematical programming for the efficient allocation of health care resources. *Journal of Health Economics* 1996;15; 641-653.

Wagstaff A. QALYs and the equity-efficiency trade-off. *Journal of Health Economics* 1991;10; 21-41.

Wagstaff A. Inequality aversion, health inequalities and health achievement. *Journal of Health Economics* 2002;21; 627-641.

Weinstein M, Zeckhauser R. Critical ratios and efficient allocation. *Journal of Public Economics* 1973;2; 147-157.

Williams A, Cookson R. 2000. Equity in health. In: Culyer AJ, Newhouse JP (Eds), *Handbook of Health Economics*. Elsevier: Amsterdam; 2000.

Wilson ECF, Rees J, Fordham RJ. Developing a prioritisation framework in an English Primary Care Trust. *Cost effectiveness and resource allocation* 2006;4.