J. N. Hooker and H. P. Williams

Combining equity and utilitarianism in a mathematical programming model

Article (Accepted version)
(Unrefereed)

Original citation:

Hooker, J. N. and Williams, H. Paul *Combining equity and utilitarianism in a mathematical programming model.* Management science. ISSN 0025-1909 (In Press)

Combining Equity and Utilitarianism in a Mathematical Programming Model

J. N. Hooker
Carnegie Mellon University
john@hooker.tepper.cmu.edu

H. P. Williams
London School of Economics
h.p.williams@lse.ac.uk

Revised 18 November 2011

Abstract

We discuss the problem of combining the conflicting objectives of equity and utilitarianism, for social policy making, in a single mathematical programming model. The definition of equity we use is the Rawlsian one of maximising the minimum utility over individuals or classes of individuals. However, when the disparity of utility becomes too great, the objective becomes progressively utilitarian. Such a model is particularly applicable to health provision but to other areas as well. Building a mixed integer/linear programming (MILP) formulation of the problem raises technical issues, as the objective function is nonconvex and the hypograph is not MILP representable in its initial form. We present a succinct formulation and show that it is "sharp" in the sense that its linear programming relaxation describes the convex hull of the feasible set (before extra resource allocation or policy constraints are added). We apply the formulation to a health care planning problem and show that instances of realistic size are easily solved by standard MILP software.

1 Introduction

The dilemma over whether to pursue policies that emphasize equity (sometimes regarded as "fairness") or utilitarianism ("total good") faces all societies. Such policies are often in conflict and have been addressed by a number of authors, such as Binmore (2005) and Sen and Williams (1982). Should we attempt to reduce differences in wealth at the expense of economic efficiency? Is equity in health provision, for example, more important than maximising the aggregate health of the population?

Utilitarianism was advocated by Bentham and Mill in the 18th and 19th centuries; that is, maximising total utility irrespective of differences between individuals or classes of individuals. Equity (egalitarianism) can be formulated in different ways. In this paper we choose the maximin principle enunciated by Rawls (1972); that is, one seeks to allocate goods so as to maximise the welfare of the worst off. There is evidence to suggest that this is considered by the majority of the population to be the most acceptable policy to pursue, at least in health matters (Yaari and Bar-Hillel, 1984). But most people regard it as unreasonable to take such a policy to its extreme; that is, to continue with such a policy when it takes too many resources from others. There is some indirect evidence for this in Williams and Cookson (2006) and Yaari and Bar-Hillel (1984). Hence we switch to a utilitarian objective in extreme circumstances.

Our discussion is perhaps most obviously relevant to health provision but is also applicable to other areas, such as facility location (Ogryczak, 1997), famine relief (Hall and Vohra, 1993), taxation (Young, 1995), communication network management (Nace and Pioro, 2008), or even to timing traffic lights, given the incompatibility between maximising traffic flow and minimising any person's maximum waiting time.

In Section 2 we propose a social welfare function, based on a suggestion of Williams and Cookson (2000), that combines equity and efficiency in the desired fashion for a two-person problem. We build a mixed integer/linear programming (MILP) formulation that maximizes the function subject to resource limitations and other constraints. This requires that some technical issues be addressed, as the function is not only nonconvex, but its hypograph is not MILP representable in its initial form.

It is not obvious how to extend the social welfare function to n persons, but we propose in Section 4 an extension that appears to capture the underlying idea. We provide a succinct MILP formulation that contains only n binary variables. We prove that despite the simplicity of the model, it is nonetheless "sharp." Jeroslow (1989) defined an MILP formulation to be sharp if its linear programming (LP) relaxation describes the (closure of) the convex hull of feasible integer solutions, making it the "best" possible formulation as a mathematical programme. Our result shows that a very compact formulation of the social welfare function can be sharp as well.

Of course, this MILP formulation only comprises the "core" of a practical model. Additional (problem-specific) constraints must be added to impose resource limitations and policy decisions, which will constrain the possible allocations of utilities. If the constraints are suitably formulated, this will result in a genuine MILP model of the allocation problem, for which integer programming methods will be required. However, the original MILP model (before adding problem-specific constraints) is the "best" possible in terms of sharpness.

Practical application often requires that resources be allocated to groups or classes rather than to individuals, where the groups may have different sizes. We show in Section 5 that the MILP formulation can be extended to this case without sacrificing sharpness. The relevant proofs appear in the Appendix. In Section 6 we apply the extended model to health care resources planning. We report computational tests indicating that the model can be solved in a few seconds even for a thousand groups or more.

There is a large literature on social welfare functions, although only a few combine equity and efficiency. The Gini coefficient, McLoone index, Atkinson's function, Hoover index, and Theil index measure inequality. These and others are discussed by Tempkin (1993). Maximin and lexicographic maximum functions aim to capture a Rawlsian fairness criterion and are discussed, for example, by Blackorby et al. (2002), Daniels (1989), Dworkin (1977), Luss (1999), Roemer (1998), and Stein (2006). Structural properties of optimal solutions for utilitarian and lexmax objectives are derived in Hooker (2010).

Nash bargaining and Raiffa-Kalai-Smorodinsky bargaining may be seen as reflecting both equity and efficiency. These and other schemes are discussed by Blackorby et al. (2002), Gaertner (2009), and Yaari and Bar-Hillel (1984). Proportional fairness objectives for communication networks are closely related to the Nash bargaining solution and are discussed by Kelly et al. (1999) and Mazumdar et al. (1991), among others. The efficiency cost of proportional and maximin fairness objectives is studied by Bertsimas et al. (to appear). Welfare functions for health care allocation are discussed by Broome (1988), Stinnett and Paltiel (1995), and Williams and Cookson (2000, 2006).

2 Problem and Basic Approach

We suppose that a population consists of individuals (or classes of individuals) and that our policies would result in an allocation of utilities $u_1, u_2, ..., u_n$ to these individuals. In the health context these utilities could be quality adjusted life years (QALYs) (Broome, 1988; Dolan, 1998).

We will endeavor to implement a policy (e.g., resource allocation) that maximises the utility of the worst off—that is, maximises $\min_i\{u_i\}$—unless this takes too many resources from the others. Following a suggestion in Williams and Cookson (2000) for the two-person case, we will switch from a Rawlsian to a utilitarian criterion when inequality exceeds a threshold; that is, when $|u_1 - u_2| \geq \Delta$. We therefore define a social welfare function that has the contours shown in Fig. 1. When $|u_1 - u_2| \leq \Delta$, the contours reflect the Rawlsian function $\min\{u_1, u_2\}$, and otherwise they reflect the utilitarian function $u_1 + u_2$. The advantage of a formulating a social welfare function is that it can be maximized, subject to resource limitations and other constraints, so as to determine the most desirable equity/efficiency tradeoff. It is not obvious how to extend this approach beyond two persons, but we will propose below an extension that captures the underlying motivation.

Maximising this social welfare function has the effect of adhering to a Rawlsian criterion unless the cost to the other party is too great. Suppose that due to limited

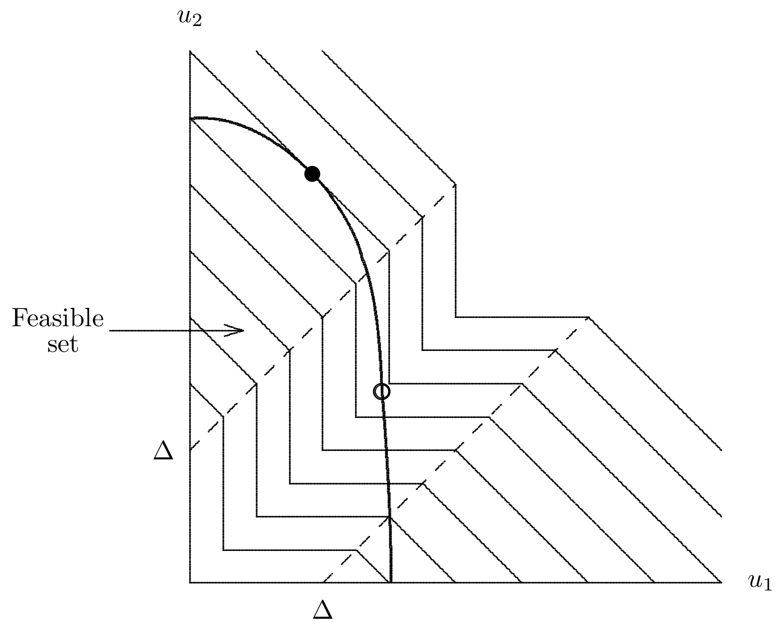

Figure 1: Contours of the social welfare function for a 2-person problem. The diagonal contours correspond to a utilitarian criterion, and the L-shaped contours to a Rawlsian criterion. The curve reflects resource limits.

health care resources, allocations of utility must lie in the region under the curve in Fig. 1. The curve indicates that person 1 is difficult to treat, because allocating all resources to person 1 is much less effective than allocating all resources to person 2. Furthermore, the marginal cost of improving person 1's health becomes very high as the limit is approached.

A purely Rawlsian solution allocates equal utility to each person, as indicated by the open circle. However, this requires great sacrifice from person 2. A small diversion of resources to person 2 would have substantial benefit while only slightly degrading the health of person 1. It may therefore be rational to switch to a utilitarian objective. In fact, the maximum value of the social welfare function, indicated by the black dot, occurs in the utilitarian region.

The level at which to set Δ is clearly judgemental and likely to be a point of disagreement among the parties concerned. However, once a value for Δ has been settled upon, maximising the social welfare function allows the same policy to be applied consistently whenever a budgeting decision is taken. It is necessary to agree on an efficiency/equity compromise only once, when the value of Δ is selected, rather than revisiting the issue every time it comes up in practice.

Furthermore, the model allows policy makers to examine the consequences of a given value of Δ across a wide variety of cases. They can compute allocations in typical scenarios for each of several values of Δ. Stakeholders can then examine each scenario and indicate which allocation they prefer. The value of Δ that results in the most popular (or least objectionable) allocation might then be selected. Once it is selected, the stakeholders can be assured that the same policy is applied consistently across the board.

One could, of course, maximize a linear combination of utilitarian and Rawlsian objectives:

$$\sum_i u_i + \alpha \min_i \{u_i\}$$

which is easy to model because it is concave. However, this raises the question of how to justify and interpret any particular multipler α. By contrast, Δ has intuitive meaning and is measured in the same units as utility. It is the level of inequality at which efficiency considerations take over. In a health care context, for example, a resource allocation in which some persons enjoy Δ QALYs more than others should begin to take efficiency into account.

3 Two-person Problem

We wish to allocate utilities u_1, u_2 to two individuals so as to maximise a social welfare function with the contours illustrated in Fig. 1. Because we want the function to be continuous, in the Rawlsian case we define its value to be $2\min\{u_1, u_2\} + \Delta$ rather than $\min\{u_1, u_2\}$. The optimization problem is therefore to maximize z subject to

$$\begin{aligned} & z \leq \begin{cases} 2\min\{u_1, u_2\} + \Delta & \text{if } |u_1 - u_2| \leq \Delta \\ u_1 + u_2 & \text{otherwise} \end{cases} \\ & u_1, u_2 \geq 0 \end{aligned} \tag{1}$$

and subject to additional constraints on u_1, u_2 that are added to represent resource limits or policy restrictions. Such constraints will be illustrated in Section 6.

We wish to write an MILP model for (1). In order for the problem to be MILP representable (Jeroslow, 1987, 1989), its hypograph must be the union of a finite number of polyhedra with the same recession directions. We do not repeat the definitions of hypograph and recession directions here but refer the reader to Jeroslow (1987, 1989), Hooker (2009), or Williams (2009). If the polyhedra do not have the same recession directions, then some innocuous constraints can be added to equalise the recession cones.

The hypograph of (1) is the union of two polyhedra, defined respectively by the two disjuncts:

$$\begin{pmatrix} z \le 2u_1 + \Delta \\ z \le 2u_2 + \Delta \\ u_1, u_2 \ge 0 \end{pmatrix} \vee \begin{pmatrix} z \le u_1 + u_2 \\ u_1, u_2 \ge 0 \end{pmatrix}$$

The first disjunct corresponds to the maximin case and the second to the utilitarian case.

The two polyhedra have different recession cones. The recession cone for the first is spanned by the four vectors

$$(u_1, u_2, z) = (1, 1, 2), (1, 0, 0), (0, 1, 0), (0, 0, -1)$$

The recession cone for the second is spanned by the vectors

$$(u_1, u_2, z) = (1, 1, 0), (1, 0, 1), (0, 0, -1)$$

However, if we add the constraints $u_1 - u_2 \le M$ and $u_2 - u_1 \le M$ to each disjunct, then the polyhedra have the same recession cone, spanned by the vectors

$$(u_1, u_2, z) = (1, 1, 2), (0, 0, -1)$$

This is illustrated in Fig. 2. The hypograph is now represented by the big-M model

$$\begin{aligned} & z \le 2u_i + \Delta + (M - \Delta)\delta, \ \ i = 1, 2 && (a) \\ & z \le u_1 + u_2 + \Delta(1 - \delta) && (b) \\ & u_1 - u_2 \le M, \ \ u_2 - u_1 \le M && (c) \\ & u_1, u_2 \ge 0, \ \ \delta \in \{0, 1\} \end{aligned} \tag{2}$$

We can also give the two polyhedra the same recession cone (namely, the origin) by imposing bounds $u_1, u_2 \le M$. In this case the formulation is the same except that constraints (c) are replaced by $u_1, u_2 \le M$.

Model (2) is a *sharp* formulation of (1), meaning that it has the tightest possible continuous relaxation. Its continuous relaxation describes a polyhedron whose projection onto the original variables is the (closure of) the convex hull of the hypograph.

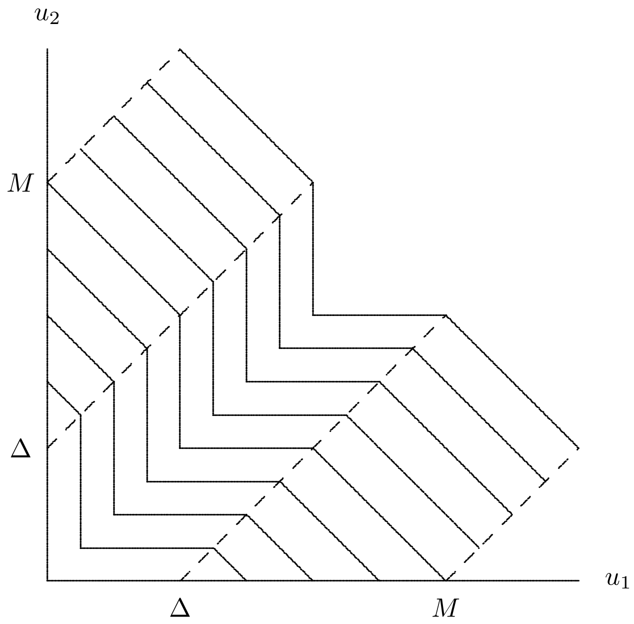

Figure 2: Modified social welfare function for which the hypograph is the union of two polyhedra with the same recession cone.

4 Many-person Problem

We now allot utilities $u_1, \ldots, u_n$ to n individuals. One way to generalize the two-person problem (1) is to observe that (1) can be written

$$\begin{aligned} &z \leq \Delta + 2u_{\min} + \max\{0, u_1 - u_{\min} - \Delta\} + \max\{0, u_2 - u_{\min} - \Delta\} \\ &u_1, u_2 \geq 0 \end{aligned} \tag{3}$$

where $u_{\min} = \min\{u_1, u_2\}$. Thus each person i makes a utilitarian contribution if u_i differs from $u_{\min}$ more than Δ. If $u_1 > u_2 + \Delta$, the first max term of (3) contributes $u_1 - u_{\min} - \Delta$ and the second max term nothing, yielding $u_1 + u_2$ altogether, and similarly if $u_2 > u_1 + \Delta$. Otherwise, both max terms vanish.

The pattern in (3) can be generalized as follows:

$$\begin{aligned} &z \leq (n-1)\Delta + nu_{\min} + \sum_{i=1}^{n} \max\{0, u_i - u_{\min} - \Delta\} \\ &u_i \geq 0, \text{ all } i \end{aligned} \tag{4}$$

where $u_{\min} = \min_i\{u_i\}$. Thus person i contributes u_i if u_i differs from $u_{\min}$ more than Δ and is otherwise represented by $u_{\min}$.

1

Suppose for illustration that $u_1 \leq \cdots \leq u_n$. If everyone makes a utilitarian contribution, then the summation in (4) becomes

$$-(n-1)\Delta - (n-1)u_1 + \sum_{i>1} u_i$$

and the inequality constraint in (4) becomes $z \leq \sum_{i=1}^n u_i$. If no two utilities differ by more than Δ, the summation in (4) vanishes, and the resulting inequality constraint yields a maximin solution. If only $u_1, \ldots, u_k$ are within Δ of u_1, the constraint becomes

$$z \leq (k-1)\Delta + ku_1 + \sum_{i=k+1}^n u_i$$

This is a combination of the lowest order statistic with the $n-k$ highest order statistics, where the lowest is multiplied by k so that all persons receive equal consideration. Thus everyone in lower stratum (within Δ of the bottom) is identified with the least advantaged person, and the lower stratum receives weight in proportion to its size.

One might achieve a roughly similar effect by giving more weight to lower terms in a linear combination of order statistics, but then the problem of justifying and interpreting weights is only compounded, because there are so many of them. Also it is difficult to model such a function for solution by optimization methods. The function proposed here requires only one parameter Δ, regardless of the number of individuals involved, and it has a practical MILP formulation.

4.1 Mixed Integer Formulation

As in the two-person case, the n-person problem can be formulated as an MILP if we suppose that

$$u_j - u_i \leq M, \quad \text{all } i, j \tag{5}$$

We can in principle write an MILP model for (4)–(5) based on a union of polyhedra similar to that used in the two-person case, but this results in exponentially many 0-1 variables. A much more compact model to maximizes z subject to

$$\begin{aligned} & z \leq (n-1)\Delta + \sum_{i=1}^n v_i && (a) \\ & u_i - \Delta \leq v_i \leq u_i - \Delta\delta_i, \ \text{all } i && (b) \\ & w \leq v_i \leq w + (M-\Delta)\delta_i, \ \text{all } i && (c) \\ & u_i \geq 0, \ \delta_i \in \{0,1\}, \ \text{all } i \end{aligned} \tag{6}$$

and again subject to resource and policy constraints. The interpretation of δ_i is that it is 0 when $u_i - u_{\min} < \Delta$ and is 1 otherwise.

Theorem 1 *The MILP model (6) is a correct formulation of problem (4)–(5).*

Proof. We must show that any feasible solution of the problem is a solution of (6), and vice-versa. To show the former, consider any feasible solution (u, z). We exhibit values of v, w, δ such that (u, z, v, w, δ) is a feasible solution of (6). Supposing without loss of generality that $u_{\min} = u_1$, set

$$w = u_1, \qquad (\delta_i, v_i) = \begin{cases} (0, u_1) & \text{if } u_i - u_1 < \Delta \\ (1, u_i - \Delta) & \text{otherwise} \end{cases} \tag{7}$$

To show that (b) and (c) in (6) are satisfied, note that when $u_i - u_1 < \Delta$, they are satisfied due to (7). When $u_i - u_1 \geq \Delta$, (b) and the first inequality in (c) are satisfied due to (7), and the second inequality in (c) is satisfied because $u_i - u_1 \leq M$ is given. To show (a), write it as

$$z \leq (n-1)\Delta + nu_1 + \sum_{\substack{i \\ u_i - u_1 < \Delta}} (v_i - u_1) + \sum_{\substack{i \\ u_i - u_1 \geq \Delta}} (v_i - u_1)$$

Substituting the values of v_i given in (7), this becomes

$$z \leq (n-1)\Delta + nu_1 + \sum_{\substack{i \\ u_i - u_1 < \Delta}} (u_1 - u_1) + \sum_{\substack{i \\ u_i - u_1 \geq \Delta}} (u_i - u_1 - \Delta)$$

which is implied by (4).

We now suppose that (u, z, v, w, δ) satisfies (6) and show that (u, z) satisfies (4) and $u_j - u_i \leq M$ for all i, j. To show the latter, note that (c) implies that $v_j - (M - \Delta)\delta_j \leq w \leq v_i$ for any i, j, and therefore

$$v_j - v_i \leq M - \Delta \tag{8}$$

But because $v_j \geq u_j - \Delta$ and $v_i \leq u_i$ due to (b), (8) implies $u_j - u_i \leq M$, as claimed. To show that (u, z) satisfies (4), write (6a) as

$$z \leq (n-1)\Delta + nu_1 + \sum_{\substack{i \\ \delta_i = 0}} (v_i - u_1) + \sum_{\substack{i \\ \delta_i = 1}} (v_i - u_1) \tag{9}$$

Each term of the first summation satisfies

$$v_i - u_1 \leq w - u_1 \leq 0 \leq (u_i - u_1 - \Delta)^+ \tag{10}$$

where the first inequality is due to (b) and $\delta_i = 0$. Noting from (b) and (c) that $w \leq v_i \leq u_i - \Delta\delta_i$ for all i, we have $w \leq u_1$, whence the second inequality in (10). Also each term of the second summation satisfies

$$v_i - u_1 \leq u_i - u_1 - \Delta \leq (u_i - u_1 - \Delta)^+$$

where the first inequality is due to (b) and $\delta_i = 1$. Inequality (9) therefore implies (4), as desired.

4.2 Proof of Sharpness

The model (6) is sharp because the projection of its continuous relaxation onto (z, u)-space is the convex hull of the original problem. The continuous relaxation of (6) is

$$\begin{aligned}
& z \leq (n-1)\Delta + \sum_{i=1}^{n} v_j && (a) \\
& u_i - \Delta \leq v_i, \;\; \text{all } i && (d_i) \\
& v_i \leq u_i - \Delta\delta_i, \;\; \text{all } i && (e_i) \\
& w \leq v_i, \;\; \text{all } i && (f_i) \\
& v_i \leq w + (M - \Delta)\delta_i, \;\; \text{all } i && (g_i) \\
& \delta_i \geq 0 \;\; \text{all } i && (h_i) \\
& \delta_i \leq 1, \;\; u_i \geq 0, \;\; \text{all } i &&
\end{aligned} \qquad (11)$$

Theorem 2 *The model (6) is a sharp formulation of the problem (4)–(5).*

Proof. The proof consists of two parts. We first show that (11) implies the following:

$$\begin{aligned}
& z \leq (n-1)\Delta + \left(1 + (n-1)\frac{\Delta}{M}\right) u_i + \left(1 - \frac{\Delta}{M}\right) \sum_{j \neq i} u_j, \text{ all } i && (k_i) \\
& u_j - u_i \leq M, \text{ all } i, j && (\ell_{ij}) \\
& u_i \geq 0, \text{ all } i &&
\end{aligned} \qquad (12)$$

We then show that every valid inequality for the original problem is implied by (12). Because (6) is a correct model of the problem, it follows that (12) describes the convex hull of the feasible set, and (6) is a sharp model.

Part I. We wish to show that (11) implies (12). We saw in the proof of Theorem 1 that (11) implies (ℓ_{ij}) for all i, j. To show that (11) implies (k_i) for any i, we show that (k_i) is a surrogate (nonnegative linear combination) of inequalities of (11). First, note that the following inequalities are surrogates of (11) for each i:

$$v_i \le \frac{\Delta}{M} w + \left(1 - \frac{\Delta}{M}\right) u_i \tag{p_i}$$

$$v_i \le u_i \tag{q_i}$$

because $(p_i) = \frac{1}{\Delta}(e_i) + \frac{1}{M-\Delta}(g_i)$ and $(q_i) = \frac{1}{\Delta}(e_i) + (h_i)$. Now we have the following for each i, j:

$$v_j \le \frac{\Delta}{M} v_i + \left(1 - \frac{\Delta}{M}\right) u_j \tag{r_{ij}}$$

because $(r_{ij}) = \frac{M}{\Delta}(p_j) + (q_i)$. Finally,

$$(k_i) = (a) + \sum_{j \neq i} (r_{ij}) + \left(1 + (n-1)\frac{\Delta}{M}\right)(q_i)$$

which shows that (k_i) is a surrogate of (11), as desired.

Part II. It remains to show that any inequality $z \le au + b$ that is valid for the problem is implied by (12). For this it is enough to show that $z \le au + b$ is dominated by a surrogate of (12).

First we observe that $(u_1, \ldots, u_n, z) = (0, \ldots, 0, (n-1)\Delta)$ is feasible in (4) and must therefore satisfy $z \le au + b$. Substituting these values into $z \le au + b$, we obtain $b \ge (n-1)\Delta$. Also, for any $t \ge 0$,

$$(u_1, \ldots, u_n, z) = (t, \ldots, t, nt + (n-1)\Delta)$$

is feasible in (4), which implies

$$\sum_i a_i \ge n - \frac{b - (n-1)\Delta}{t}$$

Letting $t \to \infty$, we get that $\sum_i a_i \ge n$. It suffices to show that any $z \le au + b$ with $\sum_i a_i = n$ is dominated by a surrogate of (12), because in this case an inequality with $\sum_i a_i > n$ can be obtained by adding multiples of $u_i \ge 0$ to an inequality with $\sum_i a_i = n$.

We let $N = \{1, \ldots, n\}$ and define index sets as follows:

$$I = \left\{ i \in N \;\middle|\; 1 - \frac{\Delta}{M} \le a_i \le 1 \right\}, \quad J = \left\{ i \in N \;\middle|\; a_i < 1 - \frac{\Delta}{M} \right\}, \quad K = N \setminus (I \cup J)$$

We next associate multipliers α_i with (k_i) and β_{ij} with (ℓ_{ij}), defined by

$$\alpha_i = \begin{cases} \dfrac{M}{n\Delta}\left(a_i - 1 + \dfrac{\Delta}{M}\right) & \text{if } i \in I \\ \dfrac{1-\alpha[I]}{n-|I|} & \text{otherwise} \end{cases}$$
$$\beta_{ij} = \begin{cases} \dfrac{1}{|K|}\left(\dfrac{n-a[I]}{n-|I|} - a_i\right) & \text{if } i \in J,\ j \in K \\ f_{ij} & \text{if } i,j \in K \text{ and } i \neq j \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where $\alpha[I] = \sum_{j\in I}\alpha_j$, and similarly for $a[I]$ and $a[K]$. The quantities f_{ij} are feasible nonnegative flows on edges (i,j) of a complete directed graph whose vertices correspond to indices in K, with a net supply of $a_i - a[K]/|K|$ at each vertex i. Such flows exist because the net supply over all vertices is $\sum_i(a_i - a[K]/|K|) = 0$.

We first show that the linear combination $\sum_i \alpha_i(k_i) + \sum_{ij}\beta_{ij}(\ell_{ij})$ is the inequality $z \le au + (n-1)\Delta$, given that $\sum_i a_i = n$. It is easily checked that $\sum_i \alpha_i = 1$, so that the linear combination has the form $z \le du + (n-1)\Delta$. It remains to show that $d = a$. We have

$$d_i = \left(1 + (n-1)\frac{\Delta}{M}\right)\alpha_i + \left(1 - \frac{\Delta}{M}\right)\sum_{j\neq i}\alpha_j + \sum_j(\beta_{ji} - \beta_{ij})$$

Using the fact that $\sum_i \alpha_i = 1$, this becomes

$$d_i = \frac{\Delta}{M}(n\alpha_i - 1) + 1 + \sum_j(\beta_{ji} - \beta_{ij}) \tag{14}$$

When $i \in I$, each $\beta_{ij} = 0$, and we immediately get from (14) that $d_i = a_i$. When $i \in J$, (14) becomes

$$d_i = \frac{n-a[I]}{n-|I|} - \sum_{j\in K}\frac{1}{|K|}\left(\frac{n-a[I]}{n-|I|} - a_i\right) = a_i$$

When $i \in K$, (14) becomes

$$\begin{aligned} d_i &= \frac{n-a[I]}{n-|I|} + \sum_{j\in J}\frac{1}{|K|}\left(\frac{n-a[I]}{n-|I|} - a_j\right) + \sum_{j\in K\setminus\{i\}}(f_{ji} - f_{ij}) \\ &= \left(1 + \frac{|J|}{|K|}\right)\frac{n-a[I]}{n-|I|} - \frac{a[J]}{|K|} + \sum_{j\in K\setminus\{i\}}(f_{ji} - f_{ij}) \end{aligned}$$

Using the fact that $a[J] = n - a[I] - a[K]$, this simplifies to

$$d_i = \frac{a[K]}{|K|} + \sum_{j \in K \setminus \{i\}} (f_{ji} - f_{ij}) \tag{15}$$

But this implies $d_i = a_i$, because the second term is the net supply at vertex i, which is $a_i - a[K]/|K|$.

We conclude that $z \leq au+(n-1)\Delta$ is the linear combination $\sum_i \alpha_i(k_i)+\sum_{ij} \beta_{ij}(\ell_{ij})$. Since $b \geq (n-1)\Delta$, $z \leq au+b$ is dominated by a surrogate of (12) and therefore implied by (6), provided we show that the multipliers α_i, β_{ij} are nonnegative.

We observe first that $\alpha_i \geq 0$ for $i \in I$ because $a_i \geq 1 - \Delta/M$, due to the definition of I. To show that $\alpha_i \geq 0$ for $i \notin I$, we note that $a_i \leq 1$ for $i \in I$ implies $\alpha_i \leq 1/n$, from the definition of α_i. Thus $\alpha[I] \leq 1$, which implies $\alpha_i \geq 0$ for $i \notin I$. To show that $\beta_{ij} \geq 0$ for $i \in J$ and $j \in K$, we note that $a_i \leq 1$ for $i \in I$ implies that $a[I] \leq |I|$, whence

$$\frac{n - a[I]}{n - |I|} \geq 1 \tag{16}$$

But $a_i < 1 - \Delta/M$ for $i \in J \setminus \{j\}$ implies $a_i \leq 1$, which along with (16) implies that $\beta_{ij} \geq 0$. Finally, $\beta_{ij} = f_{ij}$ for $i, j \in K$ is by definition a nonnegative flow.

5 Modeling Groups of Recipients

Policy makers often allocate resources to groups or classes of recipients rather than individuals. This is true in particular for health care planning, where funding for specific types of treatments is allocated to classes of patients depending on the type and prognosis of their illness. The classes generally vary in size.

In principle, such a situation can be modeled by introducing a utility variable u_i for each individual, and imposing side constraints that require individuals within a given class to receive the same allocation. However, this can result in a very large MILP model. Fortunately, it is possible to build a sharp model for the problem by allocating utility to groups rather than individuals, even when the groups have different sizes.

We therefore suppose there are m groups of recipients, and each group i has size n_i. Because each member of a group receives the same allocation, we split the utility allocated to a group evenly among the members of the group. Let u_i be the per capita utility in group i, so that the group's total utility is $n_i u_i$. The optimization problem therefore maximizes z subject to

$$
\begin{aligned}
& z \leq \left(\sum_{i=1}^{m} n_i - 1\right)\Delta + \left(\sum_{i=1}^{m} n_i\right) u_{\min} + \sum_{i=1}^{m} n_i \left(u_i - u_{\min} - \Delta\right)^{+} \\
& u_j - u_i \leq M, \;\; \text{all } i, j \\
& u_i \geq 0, \;\text{all } i
\end{aligned}
\tag{17}
$$

and subject to resource and policy constraints, where again $u_{\min} = \min_j\{u_j\}$.

The most nearly utilitarian case occurs when there is a group k that is far below the others in average utility, i.e., $u_i - u_k > \Delta$ for all $i \neq k$. In this case the inequality constraint in (17) becomes

$$z \leq (n_k - 1)\Delta + \sum_i n_i u_i$$

It is not quite utilitarian, as there is an offset that depends on which group is worst off. This is because the utilities in the worst-off group are equally low and therefore within Δ of the lowest, which means they do not receive utilitarian treatment.

5.1 The Two-group Problem

It is interesting to examine the two-group problem, which maximizes z subject to

$$
\begin{aligned}
& z \leq (n_1 + n_2 - 1)\Delta + (n_1 + n_2)u_{\min} + n_1(u_1 - u_{\min} - \Delta)^{+} + n_2(u_2 - u_{\min} - \Delta)^{+} \\
& u_1 - u_2 \leq M, \;\; u_2 - u_1 \leq M \\
& u_1, u_2 \geq 0
\end{aligned}
\tag{18}
$$

and side constraints. A graph of the model with $n_1 < n_2$ appears in Fig. 3. Note that the utilitarian contours now have slope $-n_1/n_2$ rather than -1.

5.2 MILP Model

An MILP formulation of the multi-group problem maximizes z subject to

$$
\begin{aligned}
& z \leq \left(\sum_i n_i - 1\right)\Delta + \sum_{i=1}^{n} n_i v_i && (a) \\
& u_i - \Delta \leq v_i \leq u_i - \Delta\delta_i, \;\; \text{all } i && (b) \\
& w \leq v_i \leq w + (M - \Delta)\delta_i, \;\; \text{all } i && (c) \\
& u_i \geq 0, \;\; \delta_i \in \{0, 1\}, \;\; \text{all } i &&
\end{aligned}
\tag{19}
$$

and side constraints. Again $\delta_{ij} = 1$ when $u_j - u_i \geq \Delta$.

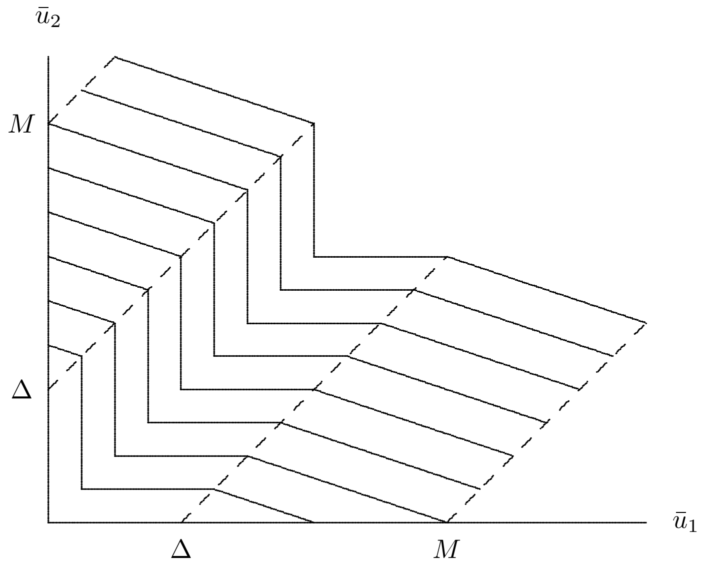

Figure 3: Contours of the social welfare function for the 2-group problem with $n_1 < n_2$.

Theorem 3 *Model (19) is a correct formulation of (17).*

The proof is similar to the proof of Theorem 1 and is given in the Appendix.

Theorem 4 *Model (19) is sharp.*

The extension to groups significantly complicates the proof of sharpness, which is likewise given in the Appendix.

6 Application to Health Care

A central problem of health care policy is to allocate scarce resources to classes of patients, depending on their prognosis and the extent to which they would benefit from various treatments. Treatments frequently have an all-or-nothing character, as in the case of a surgery, chemotherapy regimen, or organ transplant. Because utility is allocated equally to the members of a group, we must be able to model the case in which a treatment is given to all members of the group or none.

We therefore introduce a 0-1 variable y_i that is equal to 1 when everyone in group i receives a specified treatment. Let q_i be the average net gain in QALYs for a member of group i when the treatment is administered. Then the per capita utility of group i is

$$u_i = q_i y_i + \alpha_i \tag{20}$$

where α_i is the average QALYs that result from medical management without the treatment in question. If c_i is the added cost per patient of administering the treatment, the budget constraint becomes

$$\sum_i n_i c_i y_i \leq B \tag{21}$$

where B represents the funds available for providing the treatments. The resulting MILP model maximizes z subject to (19) along with constraint (20) for all i, constraint (21), and $y_i \in \{0, 1\}$ for all i.

Patients in a group may present different states of health and therefore benefit in different degrees from the same treatment. It may therefore be important in practice to divide a group into relatively homogeneous subgroups. The subgroups would be distinguished by different values of q_i and/or α_i in the above model.

To illustrate this process we present a small but fairly realistic example. The medical literature contains cost-per-QALY estimates for a wide variety of treatments, and most of the relevant papers and/or estimates are available online in the CEA Registry (Tufts Medical Center, 2010) or the United Kingdom's NICE website (NHS, 2010). However, these studies span different eras, geographical regions, and clinical settings, which result in incomparable cost estimates. In addition, most studies examine very specific treatments or interventions.

We therefore built a model around cost-per-QALY data provided by a single source, Briggs and Gray (2000), which covers a limited selection of treatments but provides estimates based on a common methodology. Briggs and Gray derive their costs per QALY in part from net QALY gains reported for these treatments by Williams (1985). In most cases, we obtained the per-patient cost c_i by multiplying the average cost per QALY in the Briggs and Gray article by the net QALY gain in the Williams article.

The data we used for the MILP model appear in Table 1. The group sizes are based approximately on various estimates of the relative frequency of each intervention in the United States. However, the relative frequency of kidney dialysis patients is reduced to one-third the prevailing rate, because otherwise this very large population would overwhelm our small example.

The groups corresponding to pacemakers, hip replacements, and aortic valve replacements are divided into three subgroups, of which subgroup B represents the average cost per QALY reported by Briggs and Gray (2000). Groups A and C reflect deviations from the average and allow policymakers to consider different prognoses among patients with the same basic disease. The nine categories of candidates for coronary artery bypass grafts (CABGs) are explicitly distinguished by Briggs and Gray, and the costs per QALY reflect their estimates.

The kidney dialysis candidates are categorized by expected lifespan while on dialysis, to reflect the fact that the cost per patient as well as the QALYs gained depend on how long the patient survives. The relative size of each category is based on survival rates for the United States reported by NKUDIC (2010). The annual cost per patient is derived from (a) Briggs and Gray's estimate of £14,000 per QALY; (b) an average of 0.688 QALYs per year of dialysis, based on converting to a 0-1 scale the Index of Well Being for such patients reported in Evans et al. (1985), which Briggs and Gray cite as their source; and (c) an average of 0.85 additional years of life obtained for each year spent on dialysis. This results in a per-capita annual dialysis cost of (14,000)(0.688)(0.85), or about £8200. Some categories are further subdivided by prognosis due to the high per-patient cost, because otherwise, funding a single category would consume a large fraction of the budget.

The expected QALYs without intervention, given by α_i, depends entirely on such population characteristics as age, general state of health, and environment. The data used here do not represent any particular population but are selected only to reflect one possible set of circumstances. The total budget is set at £3 million because this figure creates enough shortage to force some hard decisions.

Table 2 presents the results of solving the MILP model over various ranges of Δ. These results are not intended as policy recommendations, because the solution depends on population characteristics, budget, costs, and treatment options, which vary from one situation to another. Nonetheless, the results show that combining equity and efficiency in this manner can lead to interesting and perhaps unexpected conclusions.

The table shows that pacemakers are advisable under any combination of equity and efficiency, as are hip and valve replacements except in the mildest cases. However, resources shift from CABGs to kidney dialysis as Δ increases. Dialysis has a higher cost per QALY, but this is overridden by the poor prognosis without treatment when equity is more important.

There are also subtleties that one might not predict. Kidney dialysis first appears for patients with the best prognosis, for whom it is slightly less expensive per QALY, and extends to other patients as equity is more heavily emphasized. Most CABGs drop out rather suddenly when $\Delta \geq 5.6$ QALYs. In fact, CABGs for double and triple artery disease are almost always funded or defunded together, even though these subgroups have different characteristics. The same is true of dialysis decisions for most patients with less than 10 years life expectancy on dialysis. Some kidney transplants drop out when Δ reaches an intermediate value but come back in when equity dominates. A similar pattern occurs for valve replacements for mildly afflicted patients.

In general, the solution is more sensitive to Δ when Δ is between 5 and 6. This suggests that a politically acceptable compromise may place Δ in this neighborhood. It is in this range where the greatest number of interest groups are near the boundary between approval and disapproval of Δ.

As expected, the average QALYs per person generally declines as Δ increases, because larger values of Δ imply less emphasis on maximizing utility. However, due to the presence of discrete choices, there are some exceptions. The maximin solution

Table 1: Data for health care example.

Intervention	Cost per person c_i (£)	QALYs gained q_i	Cost per QALY (£)	QALYs without intervention α_i	Subgroup size n_i
Pacemaker for atrioventricular heart block					
Subgroup A	3500	3	1167	13	35
Subgroup B	3500	5	700	10	45
Subgroup C	3500	10	350	5	35
Hip replacement					
Subgroup A	3000	2	1500	3	45
Subgroup B	3000	4	750	4	45
Subgroup C	3000	8	375	5	45
Valve replacement for aortic stenosis					
Subgroup A	4500	3	1500	2.5	20
Subgroup B	4500	5	900	3	20
Subgroup C	4500	10	450	3.5	20
CABG[1] *for left main disease*					
Mild angina	3000	1.25	2400	4.75	50
Moderate angina	3000	2.25	1333	3.75	55
Severe angina	3000	2.75	1091	3.25	60
CABG for triple vessel disease					
Mild angina	3000	0.5	6000	5.5	50
Moderate angina	3000	1.25	2400	4.75	55
Severe angina	3000	2.25	1333	3.75	60
CABG for double vessel disease					
Mild angina	3000	0.25	12,000	5.75	60
Moderate angina	3000	0.75	4000	5.25	65
Severe angina	3000	1.25	2400	4.75	70
Heart transplant					
	22,500	4.5	5000	1.1	2
Kidney transplant					
Subgroup A	15,000	4	3750	1	8
Subgroup B	15,000	6	2500	1	8
Kidney dialysis					
Less than 1 year survival					
Subgroup A	5000	0.1	50,000	0.3	8
1-2 years survival					
Subgroup B	12,000	0.4	30,000	0.6	6
2-5 years survival					
Subgroup C	20,000	1.2	16,667	0.5	4
Subgroup D	28,000	1.7	16,471	0.7	4
Subgroup E	36,000	2.3	15,652	0.8	4
5-10 years survival					
Subgroup F	46,000	3.3	13,939	0.6	3
Subgroup G	56,000	3.9	14,359	0.8	2
Subgroup H	66,000	4.7	14,043	0.9	2
Subgroup I	77,000	5.4	14,259	1.1	2
At least 10 years survival					
Subgroup J	88,000	6.5	13,538	0.9	2
Subgroup K	100,000	7.4	13,514	1.0	1
Subgroup L	111,000	8.2	13,537	1.2	1

[1]Coronary artery bypass graft

Table 2: Results of health care example. A 1 in the body of the table indicates that the treatment is given to all members of a subgroup, and 0 that it is given to none. The subgroups are defined in Table 1. The last column indicates the average expected QALYs per person. These results are not intended as general policy recommendations, because they reflect one particular set of population characteristics, costs, and treatment options, which may differ substantially from one situation to another.

Δ range	Pace-maker	Hip repl.	Aortic valve	CABG L	CABG 3	CABG 2	Heart trans.	Kidney trans.	Kidney dialysis < 1	1-2	2-5	5-10	> 10	Avg. QALYs
0–3.3	111	111	111	111	111	111	1	11	0	0	000	0000	000	7.54
3.4–4.0	111	111	111	111	111	111	0	11	1	0	000	0000	000	7.54
4.0–4.4	111	111	111	111	111	111	0	01	1	0	000	0000	001	7.51
4.5–5.01	111	011	111	111	111	111	1	01	1	0	000	0000	011	7.43
5.02–5.55	111	011	011	111	111	111	0	01	1	0	000	0001	011	7.36
5.56–5.58	111	011	011	111	111	011	0	01	1	0	000	0001	111	7.36
5.59	111	011	011	110	111	111	0	01	1	0	000	0001	111	7.20
5.60–13.1	111	111	111	101	000	000	1	11	1	0	111	1111	111	7.06
13.2–14.2	111	011	111	011	000	000	1	11	1	1	111	1111	111	7.03
14.3–15.4	111	111	111	011	000	000	1	11	1	1	101	1111	111	7.13
15.5–up	111	011	111	011	001	000	1	11	1	0	011	1111	111	7.19

Table 3: Solution times in seconds for m groups and different values of Δ. Instances with more than a few hundred groups seem very unlikely to occur in practice.

m	Δ = 0	1	2	3	4	5	6	∞
330	0.02	1.2	0.67	0.56	0.50	0.30	0.03	0.02
660	0.03	4.1	1.6	1.6	0.92	0.80	0.05	0.02
990	0.02	5.2	3.1	3.6	1.5	1.5	0.08	0.02
1320	0.00	15	4.3	4.2	2.7	3.0	0.09	0.02
1980	0.02	24	11	11	11	5.4	0.14	0.02
2640	0.00	32	19	14	8.6	8.8	0.19	0.02
3300	0.17	51	43	44	34	13	0.25	0.02

($\Delta \geq 15.5$) results in greater utility than solutions corresponding to $5.60 \leq \Delta \leq 15.4$. One might argue that solutions in this range should be eliminated because they are dominated by the maximin solution with respect to both utility and equity.

This small problem, which allocates utilities u_i to 33 groups and contains 1089 integer variables, was solved in a small fraction of a second. We created much larger instances by making k copies of each group and increasing the budget by a factor of k, for $k = 10, 20, 30, 40, 60, 80, 100$. We solved the instances using CPLEX 12.2 on a desktop PC running Windows XP with a Pentium 2.8 GHz dual-core processor. The computation times appear in Table 3.

Interestingly, the problem is harder to solve for intermediate values of Δ than for the pure utilitarian and Rawlsian cases. Nonetheless, it is readily solved for any value of Δ, even when there are upwards of 3000 groups. This may be due in part to the sharp model of the social welfare function.

Models with hundreds or thousands of groups are probably too large to interpret in any case, due to the complexity of interactions. A more practical approach is to identify treatments worthy of funding for any reasonable Δ, based on a first-cut model with broad treatment categories. These treatments can be fully funded and removed from the problem, allowing policy analysts to subdivide the more controversial categories for closer scrutiny while keeping the model size within bounds. Solution times for such models will be negligible.

7 Conclusions

We showed how to formulate a social welfare function that combines equity and efficiency in a fashion that often seems reasonable, particularly in a health care context. It captures the idea that the worst-off should receive highest priority until this requires too much sacrifice from others. The threshold is reflected by a single parameter Δ that measures the level of inequality at which a utilitarian objective begins to take over from a Rawlsian objective.

We proposed what seems to be a natural generalization of the social welfare function to the n-person case. Although formulating an MILP model of the problem raises technical issues, we provided a compact MILP formulation with only n binary variables. We proved that, despite its simplicity, the model is sharp and therefore provides the best possible linear relaxation of the social welfare function. We also showed that it can be extended to groups of individuals without sacrificing sharpness. Finally, we illustrated how to adapt the extended model to a realistic health care problem and showed that life-sized instances can be easily solved using widely available MILP software.

Variants of our model are clearly possible. For example:

(i) Instead of working with a fixed Δ, we could allow it to vary with the magnitude of the values of the utilities u_j.

(ii) We could combine a utilitarian objective with a lexicographic maximum rather than a Rawlsian maximin objective.

A more ambitious but essential research goal is to find a way to justify a choice of Δ on principle rather than by political compromise. This is a task for philosophical as well as mathematical analysis.

8 Acknowledgements

We thank Mara Airoldi, Nikos Argyris, and Alec Morton of the London School of Economics for their advice and assistance in searching health policy literature. We also thank anonymous reviewers for suggestions that led to significant improvements in the paper, including a simplification of the model.

Appendix

Proof of Theorem 3. We show that any feasible solution of problem (17) is a solution of (19), and vice-versa. To show the former, consider any feasible solution (u, z). We exhibit values of v, w, δ such that (u, z, v, w, δ) is a feasible solution of (19). Supposing without loss of generality that $u_{\min} = u_1$, let (v, w, δ) be as in (7). It can be shown as in the proof of Theorem 1 that (b) and (c) in (19) are satisfied. To show (a), write it as

$$z \leq \left(\sum_i n_i - 1 \right) \Delta + \left(\sum_i n_i \right) u_1 + \sum_{\substack{i \\ u_i - u_1 < \Delta}} n_i(v_i - u_1) + \sum_{\substack{i \\ u_i - u_1 \geq \Delta}} n_i(v_i - u_1)$$

Substituting the values of v_i given in (7), this becomes

$$z \leq \left(\sum_i n_i - 1 \right) \Delta + \left(\sum_i n_i \right) u_1 + \sum_{\substack{i \\ u_i - u_1 < \Delta}} n_i(u_1 - u_1) + \sum_{\substack{i \\ u_i - u_1 \geq \Delta}} n_i(u_i - u_1 - \Delta)$$

which is implied by (17).

We now suppose that (u, z, v, w, δ) is feasible in (19) and show that (u, z) is feasible in (17). It can be demonstrated as in the proof of Theorem 1 that u satisfies the second constraint of (17). To show that (u, z) satisfies the first constraint, write (19a) as

$$z \leq \left(\sum_i n_i - 1 \right) \Delta + \left(\sum_i n_i \right) u_1 + \sum_{\substack{i \\ \delta_i = 0}} n_i(v_i - u_1) + \sum_{\substack{i \\ \delta_i = 1}} n_i(v_i - u_1) \qquad (22)$$

Each term of the first summation satisfies

$$n_i(v_i - u_1) \leq n_i(w - u_1) \leq 0 \leq n_i(u_i - u_1 - \Delta)^+ \qquad (23)$$

where the first inequality is due to (b) and $\delta_i = 0$. Noting from (b) and (c) that $w \leq v_i \leq u_i - \Delta\delta_i$ for all i, we have $w \leq u_1$, whence the second inequality in (23). Also each term of the second summation satisfies

$$n_i(v_i - u_1) \leq n_i(u_i - u_1 - \Delta) \leq n_i(u_i - u_1 - \Delta)^+$$

where the first inequality is due to (b) and $\delta_i = 1$. Inequality (22) therefore implies the first constraint of (17), as desired.

Proof of Theorem 4. The proof has two parts. We first show that the continuous relaxation of (19) implies the following:

$$
\begin{aligned}
& z \leq \left(\sum_j n_j - 1 \right) \Delta + \left(n_i + \frac{\Delta}{M} \sum_{j \neq i} n_j \right) u_i + \left(1 - \frac{\Delta}{M} \right) \sum_{j \neq i} n_j u_j, \text{ all } i && (k_i) \\
& u_j - u_i \leq M, \text{ all } i, j && (\ell_{ij}) \\
& u_i \geq 0, \text{ all } i
\end{aligned}
\tag{24}
$$

We then show that every valid inequality for the orginal problem (17) is implied by (24). Because (19) is a correct model of the problem, it follows that (24) describes the convex hull of the feasible set, and (19) is a sharp model.

Part I. We wish to show that the continuous relaxation of (19) implies (24). We saw in the proof of Theorem 3 that it implies (ℓ_{ij}) for all i, j. To show that the continuous relaxation of (19) implies (k_i) for any i, we show that (k_i) is a surrogate of the relaxation. The following surrogates are derived in the proof of Theorem 2:

$$
\begin{aligned}
& v_i \leq u_i && (q_i) \\
& v_j \leq \frac{\Delta}{M} v_i + \left(1 - \frac{\Delta}{M} \right) u_j && (r_{ij})
\end{aligned}
$$

Now

$$
(k_i) = (a) + \sum_{j \neq i} n_j (r_{ij}) + \left(n_i + \frac{\Delta}{M} \sum_{j \neq i} n_j \right) (q_i)
$$

which shows that (k_i) is a surrogate, as desired.

Part II. It remains to show that any inequality $z \leq au + b$ that is valid for the problem is implied by (24). For this it is enough to show that $z \leq au + b$ is dominated by a surrogate of (24).

First we observe that $(u_1, \ldots, u_n, z) = (0, \ldots, 0, (n[N] - 1)\Delta)$ is feasible in (17) and must therefore satisfy $z \leq au + b$. Substituting these values into $z \leq au + b$, we obtain $b \geq (n[N] - 1)\Delta$. Also, for any $t \geq 0$,

$$
(u_1, \ldots, u_n, z) = (t, \ \ldots, \ t, \ tn[N] + (n[N] - 1)\, \Delta)
$$

is feasible in (17), which implies

$$
a[N] \geq n[N] - \frac{b - (n[N] - 1)\, \Delta}{t}
$$

Letting $t \to \infty$, we get that $a[N] \geq n[N]$. It suffices to show that any $z \leq au + b$ with $a[N] = n[N]$ is dominated by a surrogate of (24), because in this case an inequality with $a[N] > n[N]$ can be obtained by adding multiples of $u_i \geq 0$ to an inequality with $a[N] = n[N]$.

We define index sets as follows:

$$\begin{aligned} I &= \left\{ i \in N \;\middle|\; n_i \left(1 - \frac{\Delta}{M}\right) \leq a_i \leq n_i \right\} \\ J &= \left\{ i \in N \;\middle|\; a_i < n_i \left(1 - \frac{\Delta}{M}\right) \right\}, \\ K &= N \setminus (I \cup J) \end{aligned}$$

We next associate multipliers with (24) as shown and define them as follows:

$$\begin{aligned} \alpha_i &= \begin{cases} \dfrac{1}{n[N]} \dfrac{M}{\Delta} \left(a_i - n_i \left(1 - \dfrac{\Delta}{M}\right)\right) & \text{if } i \in I \\ \dfrac{1 - \alpha[I]}{m - |I|} & \text{otherwise} \end{cases} \\ \beta_{ij} &= \begin{cases} \dfrac{1}{|K|} (S_i - a_i) & \text{if } i \in J,\, j \in K \\ f_{ij} & \text{if } i, j \in K \text{ and } i \neq j \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{25}$$

where

$$S_i = \left(1 - \frac{\Delta}{M}\right) n_i + \frac{n[N] - (n[N] - n[I]) \left(1 - \frac{\Delta}{M}\right) - a[I]}{m - |I|}$$

The quantities f_{ij} are feasible nonnegative flows on edges (i, j) of a complete directed graph whose vertices correspond to indices in K, with a net supply of

$$s_i = a_i - \left(n_i - \frac{n[K]}{|K|}\right)\left(1 - \frac{\Delta}{M}\right) - \frac{a[K]}{|K|}$$

at each vertex i. Such flows exist because the net supply over all vertices is $\sum_{i \in K} s_i = 0$.

We first show that the linear combination $\sum_i \alpha_i (k_i) + \sum_{ij} \beta_{ij} \ell_{ij}$ is the inequality $z \leq au + (n[N] - 1)\Delta$, given that $\sum_i a_i = n[N]$. It is easily checked that $\sum_i \alpha_i = 1$, so that the linear combination has the form $z \leq du + (n[N] - 1)\Delta$. It remains to show that $d = a$. We have

$$d_i = \left(n_i + \frac{\Delta}{M} \sum_{j \neq i} n_j\right) \alpha_i + n_i \left(1 - \frac{\Delta}{M}\right) \sum_{j \neq i} \alpha_j + \sum_{j \neq i} (\beta_{ji} - \beta_{ij}) \tag{26}$$

Using the fact that $\sum_j \alpha_j = 1$, this implies

$$d_i = \frac{\Delta}{M} n[N]\alpha_i + n_i\left(1 - \frac{\Delta}{M}\right) + \sum_{j \neq i} (\beta_{ji} - \beta_{ij}) \tag{27}$$

When $i \in I$, each $\beta_{ij} = 0$, and we get from (27) that $d_i = a_i$. When $i \in J$, (27) becomes

$$d_i = S_i - \sum_{j \in K} \frac{1}{|K|}(S_i - a_i) = a_i$$

When $i \in K$, (27) becomes

$$\begin{aligned} d_i &= S_i + \sum_{j \in J} \frac{1}{|K|}(S_j - a_j) + \sum_{j \in K \setminus \{i\}} (f_{ji} - f_{ij}) \\ &= \frac{|K|S_i + S[K] - a[J]}{|K|} + \sum_{j \in K \setminus \{i\}} (f_{ji} - f_{ij}) \end{aligned}$$

Using the definition of S_i and the fact that $n[N] = a[I] + a[J] + a[K]$, this becomes

$$d_i = \left(1 - \frac{\Delta}{M}\right) n_i - \frac{1}{|K|}\left((n[N] - n[I] - n[J])\left(1 - \frac{\Delta}{M}\right) - a[K]\right) + \sum_{j \in K \setminus \{i\}} (f_{ji} - f_{ij})$$

Using the fact that $n[N] = n[I] + n[J] + n[K]$, this simplifies to

$$\begin{aligned} d_i &= \left(1 - \frac{\Delta}{M}\right)\left(n_i - \frac{n[K]}{|K|}\right) + \frac{a[K]}{|K|} + \sum_{j \in K \setminus \{i\}} (f_{ji} - f_{ij}) \\ &= a_i - s_i + \sum_{j \in K \setminus \{i\}} (f_{ji} - f_{ij}) \end{aligned} \tag{28}$$

Because the summation is just the net supply s_i at node i, this yields $d_i = a_i$, as desired.

We conclude that $z \leq au + (n[N] - 1)\Delta$ is the linear combination $\sum_i \alpha_i(k_i) + \sum_{ij} \beta_{ij}\ell_{ij}$. Since $b \geq (n[N] - 1)\Delta$, $z \leq au + b$ is dominated by a surrogate of (24) and therefore implied by (19), provided we show that the multipliers in (25) are nonnegative.

We observe first that $\alpha_i \geq 0$ for $i \in I$ because $a_i \geq n_i(1 - \Delta/M)$, due to the definition of I. To show that $\alpha_i \geq 0$ for $i \notin I$, we note that $a_i \leq n_i$ for $i \in I$ implies $\alpha_i \leq 1/n[N]$, from the definition of α_i. Thus $\alpha(I) \leq 1$, which implies $\alpha_i \geq 0$ for $i \notin I$. To show that $\beta_{ij} \geq 0$ for $i \in J$ and $j \in K$, we note that $a_i \leq n_i$ for $i \in I$ implies that $a[I] \leq n[I]$, whence

$$S_i \geq \left(1 - \frac{\Delta}{M}\right) n_i + \frac{n[N] - n[I]}{m - |I|} \geq a_i + \frac{n[N] - n[I]}{m - |I|} \geq a_i \tag{29}$$

where the second inequality is due to the fact that $a_i < (1 - \Delta/M)n_i$ for $i \in J$. But (29) and the definition of β_{ij} imply that $\beta_{ij} \geq 0$ for $i \in J$ and $j \in K$. Finally, $\beta_{ij} = f_{ij}$ for $i, j \in K$ is by definition a nonnegative flow.

References

D. Bertsimas, V. F. Farias, and N. Trichakis. The price of fairness. *Operations Research*, to appear.

K. Binmore. *Natural Justice.* Oxford University Press, 2005.

C. Blackorby, W. Bossert, and D. Donaldson. Utilitarianism and the theory of justice. In K. Arrow, A. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare, Vol. 1*, volume 19 of *Handbooks in Economics*, pages 543–596. Elsevier, Amsterdam, 2002.

A. Briggs and A. Gray. Using cost effectiveness information. *British Medical Journal*, 320:246, 2000.

J. Broome. Fairness and QALYs. *Philosophy and Medical Welfare*, 3:57–73, 1988.

N. Daniels. *Reading Rawls: Critical Studies on Rawls' "A Theory of Justice"*. Stanford University Press, 1989.

P. Dolan. The measurement of individual utility and social welfare. *Journal of Health Economics*, 17:39–52, 1998.

R. Dworkin. *Taking Rights Seriously.* Harvard University Press, Cambridge, MA, 1977.

R. W. Evans, D. L. Manninen, L. P. Garrison, L. G. Hart, C. R. Blagg, R. A. Gutman, A. R. Hill, and E. G. Lowrie. The quality of life of patients with end-stage renal disease. *New England Journal of Medicine*, 312:553–559, 1985.

W. Gaertner. *A Primer in Social Choice Theory.* Oxford University Press, Cambridge, MA, 2009.

N. G. Hall and R. V. Vohra. Towards equitable distribution via proportional equity constraints. *Mathematical Programming*, 58:287–294, 1993.

J. N. Hooker. A principled approach to mixed integer/linear problem formulation. In J. W. Chinneck, B. Kristjansson, and M. Saltzman, editors, *Operations Research and Cyber-Infrastructure (ICS 2009 Proceedings)*, pages 79–100. Springer, 2009.

J. N. Hooker. Optimality conditions for distributive justice. *International Transactions in Operational Research*, 17:485–505, 2010.

R. G. Jeroslow. Representability in mixed integer programming, I: Characterization results. *Discrete Applied Mathematics*, 17:223–243, 1987.

R. G. Jeroslow. *Mixed Integer Formulation*, volume 40 of *Logic-Based Decision Support*. North-Holland, Amsterdam, 1989.

F. P. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate control in communication networks: Shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49:237–252, 1999.

H. Luss. On equitable resource allocation problems: A lexicographic minimax approach. *Operations Research*, 47:361–378, 1999.

R. Mazumdar, L. Mason, and C. Douligeris. Fairness in network optimal flow control: Optimality of product forms. *IEEE Transactions on Communications*, 39:775–782, 1991.

D. Nace and M. Pioro. Max-min fairness and its applications to routing and load-balancing in communication networks: A tutorial. *IEEE Communications Surveys & Tutorials*, 10:5–17, 2008.

NHS. National Institute for Health and Clinical Excellence (NICE) Website, National Health Service, United Kingdom, 2010. `http://www.nice.org.uk/`.

NKUDIC. Kidney and Urologic Diseases Statistics for the United States. National Kidney and Urologic Diseases Information Clearing House, 2010. `http://kidney.niddk.nih.gov/kudiseases/pubs/kustats/`.

W. Ogryczak. On the lexicographic minimax approach to location problems. *European Journal of Operational Research*, 100:218–223, 1997.

J. Rawls. *A Theory of Justice*. Oxford University Press, 1972.

J. E. Roemer. *Theories of Distributive Justice*. Harvard University Press, Cambridge, MA, 1998.

A. Sen and B. Williams, editors. *Utilitarianism and Beyond*. Cambridge University Press, 1982.

M. S. Stein. *Distributive Justice and Disability: Utilitarianism against Egalitarianism*. Stanford University Press, 2006.

A. A. Stinnett and A. D. Paltiel. Mathematical programming for the efficient allocation of health resources. *Journal of Health*, 15:641–653, 1995.

L. S. Tempkin. *Inequality.* Oxford University Press, 1993.

Tufts Medical Center. *CEA Registry.* Center for the Evaluation of Value and Risk in Health, Institute for Clinical Research and Health Policy Studies, 2010. `https://research.tufts-nemc.org/cear/`.

A. Williams. Economics of coronary artery bypass grafting. *British Medical Journal*, 291:326–329, 1985.

A. Williams and R. Cookson. Equity in health. In A. J. Culyer and J. P. Newhouse, editors, *Handbook of Health Economics.* Elsevier Science, 2000.

A. Williams and R. Cookson. Equity-efficiency trade-offs in health technology assessment. *International Journal of Technology Assessment in Health Care*, 22:1–9, 2006.

H. P. Williams. *Logic and Integer Programming.* Springer, Berlin, 2009.

M. E. Yaari and M. Bar-Hillel. On dividing justly. *Social Choice and Welfare*, 1:1–24, 1984.

H. P. Young. *Equity in Theory and Practice.* Princeton University Press, 1995.