ELSEVIER Journal of Health Economics 17 (1998) 39–52

JOURNAL OF HEALTH ECONOMICS

The measurement of individual utility and social welfare

Paul Dolan *

Department of Economics, University of Newcastle, Newcastle-upon-Tyne NE1 7RU, UK

Received 1 November 1994; revised 1 February 1997

Abstract

It has been suggested by a number of economists that decisions about how to allocate scarce health care resources should be informed by the cost per quality-adjusted life-years (QALYs) of the different alternatives. One of the criticisms of the QALY approach is that it is based on the measurement of individual utility; yet the values elicited are used to inform social choice. In this respect, it is argued that the QALY approach fails to take account of distributional issues that are known to be important in the context of health care. This paper addresses this issue and presents an approach grounded in microeconomic theory that is flexible enough to deal with a wide range of efficiency–equity trade-offs, while making the nature of the trade-off transparent. In addition, it is an approach that is relatively simple to investigate empirically, and the results of a preliminary study are presented as illustration of this. © 1998 Elsevier Science B.V.

JEL classification: I10

Keywords: Utility; Social welfare; Equity; Indifference curves; Person trade-off; QALYs

1. Introduction

The quality-adjusted life-year (QALY) approach to health care evaluation attempts to combine the value of quality-of-life with the value of length of life into a single index number, which may then be used as a currency in which the benefits of health care interventions can be expressed. In the simplest case, in which a

* Corresponding author. Fax: +44-0-191-222-6548; e-mail: p.h.r.dolan@ncl.ac.uk.

0167-6296/98/$19.00
PII S0167-6296(97)00022-2

person remains in the same health state for a number of years, QALYs (assuming no discounting) are calculated according to the following formula:

$$\text{QALY} = H \times Y$$

where H is the relative weight attached to a particular health state and Y is the number of years spent in that health state. [1] Although QALYs can be used to measure the benefit derived from different therapies by an individual patient, in this paper they are discussed in terms of their use in the allocation of scarce health care resources among different patients.

Nord (1994) suggests that although the number of QALYs gained from different interventions only represent the (unweighted) sum of gains in individual utility, they have been used to represent social value, whereby they are seen as a measure of society's preferences over different health care outcomes. Elsewhere, he cites the seminal paper by Weinstein and Stason (1977) as evidence that valuations have been used in this wider sense: "Alternative programs or services are then ranked, from the lowest value to the highest, and selected from the top until available resources are exhausted" (Nord et al., 1993).

According to Weinstein and Stason this ranking of alternatives takes place according to the aggregate unweighted number of QALYs gained, i.e., those that yield more QALYs are ranked higher than those that yield less. This defines the objectives of the health care system in terms of the maximisation of health gain, and is consistent with defining need in terms of capacity to benefit: that Weinstein and Stason and Williams define need in this way is therefore not surprising. In this paper, an allocation of resources that maximises the number of QALYs gained is defined as an efficient one. Of course, it is possible that people would want decision-makers, when choosing between alternatives, to be also concerned with how those QALYs are distributed; and again, different definitions of need are relevant. For example, if need is defined in terms of ill health—those in the worst health states are those most in need of treatment—then pre-treatment health status becomes the most important consideration in determining priorities. Alternatively, if need is defined in terms of final health status, then post-treatment health status is more important. There are clearly a number of other definitions of need (for a more detailed discussion see Culyer, 1995), each with different implications for the allocation of resources, but the definitions cited above highlight the tension between efficiency (defined in terms of health gain) on the one hand, and concerns for equity (defined in terms of pre- and post-treatment health status) on the other.

[1] When H changes over time, the QALY algorithm assumes that the utility derived from the whole profile is equal to the sum of the QALYs derived from each health state. In other words, it is assumed that each individual's utility function is strongly separable on the time dimension, i.e., $U(H^1, H^2, \ldots H^n; Y^1, Y^2, \ldots Y^n) = U(H^1) \times Y^1 + U(H^2) \times Y^2 + \ldots + U(H^n) \times Y^n$. While recognising that this assumption is a restrictive one, issues regarding its appropriateness are not addressed in this paper.

It is likely that people would want resource allocation decisions to be informed both by efficiency and equity considerations. A number of authors have addressed this possibility in terms of an efficiency–equity trade-off (Wagstaff, 1991; Mooney and Olsen, 1991) and, in drawing a distinction between QALYs as a measure of individual utility, on the one hand, and as a measure of social value, on the other, Nord focuses our attention directly on this trade-off.

If health state utilities [2] are interpreted as measures of social value, then their appropriateness can be tested by asking respondents whether they agree with the consequences in terms of the implied priorities for health care. Following Rawls (1971), Nord refers to this as a test of reflective equilibrium. To address this issue when determining the weights that are attached to different health states, Nord suggests that we move away from methods based on the assessment of individual utility, such as the standard gamble (SG) or the time trade-off (TTO), towards methods based on the assessment of social preferences, such as the person trade-off (PTO). Using this approach, respondents indicate the number of people in one health state they would need to be able to treat (with a specified outcome) to make them indifferent to treating a given number in another health state (again with a specified outcome). Valuations from this technique can be seen as representing the trade-offs that people are prepared to make between gains in (length and) quality of life, and number of persons treated.

However, Nord's proposal is problematic. Responses to PTO questions contain the relative weights a respondent attaches to at least four things: (1) the severity of the pre-intervention health state; (2) the severity of the post-intervention health state; (3) the health gain as a result of intervening; and (4) the number of persons treated. It is impossible from answers to PTO questions to disentangle what are the relative weights attached to each of these considerations (and consequently what the most appropriate definition(s) of need is/are). While all four are likely to be important, different weights attached to each may have quite different implications for resource allocation decisions. In addition, the fact that respondents are asked to weigh up a number of quite diverse things when thinking about their answers to PTO questions, increases the likelihood of cognitive overload, and thus may reduce the validity of responses.

In this paper, an alternative approach is suggested, using a particular class of health-related social welfare functions (HRSWF) [3], which allows efficiency and equity to be considered independently. A particular functional form is postulated,

[2] A distinction is often made between utility, a measure of preferences under uncertainty, and value, a measure of preferences under certainty (for example, Dyer and Sarin, 1982). Thus, standard gamble scores are often referred to as utilities, TTO ones as values. However, in this paper the term 'utility' is used in both cases.

[3] Although many authors define the social welfare function over the utility space generally, because the concern here is with health-related utility, for consistency the term 'health-related' social welfare function will be used throughout this paper.

which is sufficiently flexible to represent a wide range of social preferences. More importantly, the framework suggested allows a number of different hypotheses to be tested in a relatively straightforward way.

2. The social welfare function

Following Ng (1983), we may characterise social welfare by a vector of individual welfares, $(W^1, W^2, \ldots, W^I)$, where W^i is the welfare (or 'good') of the ith individual and I is the number of individuals. Economists have typically argued that individuals are the best judges of their own well-being, and that social welfare depends only on the welfare of persons in society. [4] A Bergsonian HRSWF (Bergson, 1938) may then be written as

$$W = f(W^1, W^2, \ldots, W^I) \tag{1}$$

where the precise form of f is unspecified other than that it is strictly increasing in all of its arguments. In this way, welfare economics can be said to be written largely from a consequentialist and individualistic standpoint, implying a refusal to adopt a paternalistic attitude. [5]

According to the Pareto criterion (Pareto, 1935), an increase in some W^i and decrease in no W^i is a sufficient condition for an increase in social welfare. [6] Therefore,

$$\frac{\delta f}{\delta W^i} \geq 0 \text{ for all } i \tag{2}$$

Another more restrictive definition of social welfare is the (Benthamite) utilitarian concept of the sum of individual happiness

$$W = U^1 + U^2 + \ldots U^I = \sum_i U^i \tag{3}$$

[4] These value judgements are of very general appeal and hence it is sometimes claimed that implications drawn from them will be uncontroversial. However, there are conditions (for example, concerning the consumption of hard drugs) under which many people might think it justifiable to override individual preferences.

[5] Paternalism is acceptable under the social decision-making approach to resource allocation, which suggests that policies that maximise the objectives of the decision-making unit (for example, the NHS) should be adopted (for more details of this paradigm, see Sugden and Williams, 1978). In this paper, however, it is not only assumed that individuals are the best judges of their own (real or hypothetical) health state, but also that the objectives of government are defined in terms of individual preferences, i.e., in terms of arguments in the SWF.

[6] Some economists argue that this is a sufficient and a necessary condition for an improvement in social welfare (see, for example, Gravelle and Rees, 1981). However, this very narrow interpretation implies that improvements in social welfare can only be brought about from a policy change in the (highly unlikely) event that nobody loses (and at least one person gains) from the proposed change. Thus, others (for example, Sugden, 1981) suggest that the Pareto criterion is only sufficient for an improvement in social welfare.

where U^i is a utility index representing the preferences of individual i.[7] The advantage of this approach is that the HRSWF aggregates individual utilities in a direct and transparent manner.

In the discussion that follows, it is necessary to assume that it is possible to make interpersonal comparisons of utility. It is now well established that different HRSWFs require different types of comparability (see Sen, 1977). For example, maximising the sum of individual utilities requires that differences in utilities can be compared (referred to as unit comparability) while adoption of the Rawlsian criterion of maximising the welfare of the worst-off individual requires only that we know whether one person is better or worse off than another (referred to as level comparability). For the purpose of this paper, full comparability, which subsumes both level and unit comparability, is required. By going beyond individual orderings, problems associated with Arrow's General Possibility Theorem (Arrow, 1951) are avoided.

The application of a utilitarian HRSWF to health care implies that HRSW is maximised when the total number of QALYs gained (subject to a budget constraint) is maximised, irrespective of how those QALYs are distributed. It is this approach, Nord et al. (1993) appear to object to, claiming that "The rule is almost certainly defective as it ignores distributional considerations and issues of entitlement that are known to be of importance in decision-making, especially in the health sector". I agree. But the utilitarian approach is only *one* approach to deriving a HRSWF from individual utilities. Another might be to adopt a decision rule that gives greater weight to one individual's utility than to another's. For example, a Rawlsian 'maximin' approach would require giving greatest weight to the treatment of the most seriously ill individual.

Therefore, taking account of equity and distributive considerations is not inconsistent with the measurement of individual utility, nor is it inconsistent with the interpersonal comparison of individual utilities. For example, the conclusion that individual i is in better health than individual j, and that i's gain in health from a change from Y to Z will exceed j's loss (i.e., a positive statement [8]) does not imply what ought to be done until an objective function (i.e., a normative statement) is specified. Therefore, if our objective is to maximise the sum of health utilities (i.e., the utilitarian HRSWF), we choose Z; if we want to maximise the health of the worst-off individual (i.e., the Rawlsian HRSWF), we choose Y. Of course, there are a number of other objectives that we may wish to satisfy, and these will be considered below.

The measurement of individual utilities, then, provides us with the flexibility to formulate (and even subsequently revise) any number of possible HRSWFs from

[7] In this paper, it is assumed that U_i is a cardinal utility index, unique up to a positive proportionate transformation, i.e., if U_i is a representation of the individual's preferences then the only admissible transformation is aU_i where $a > 0$.

[8] This assumes that an agreed unit of value applicable to each individual has been established.

them. The direct measurement of social value (as in the PTO) is far less flexible in that only one HRSWF can be formulated; a HRSWF for which it is impossible to disentangle the relative contributions of potentially numerous efficiency and equity considerations. Moreover, the process of collapsing individual utilities into an overall HRSWF makes explicit the assumptions and philosophical basis on which such aggregation is based. This means that the precise form of the HRSWF is open to debate, which is more likely to result in it being 'correctly' specified.

3. Distributive justice and the social welfare function

If we assume that HRSW is a function of individual utilities [9], we can use the tools of welfare economics to represent a number of different HRSWFs. In Fig. 1, the utility derived from different states of health by two (or two groups of) individuals, i and j, are shown on the x and y axes, respectively. [10] Four HRSWFs are postulated: (1) The utilitarian HRSWF (UB): a straight line drawn at right angles to the 45° line, indicating that maximising total health gain is the objective, irrespective of distributional considerations. (2) The Rawlsian HRSWF (UR): welfare is not increased unless the health state of the most seriously ill individual is improved. (3) The convex HRSWF (UC): implies that there exists a trade-off between efficiency (i.e., maximising health) and equity (i.e., greater concern for those in poorer health). (4) The concave HRSWF (UI): implies inequality proneness. [11]

All these possible formulations can be represented by a class of HRSWF recently employed by Jones-Lee and Loomes (1995), which is itself a variant of a class first proposed by Atkinson (1970).

$$W=\frac{1}{A}\left[(u_i)^A+B(u_j)^A\right],\ A\neq 0, 0<B\leq 1 \qquad (4)$$

$$W=\ln(u_i)+B\ln(u_j),\ A=0, 0<B\leq 1 \qquad (5)$$

The parameter A determines the curvature of the iso-welfare loci, thereby reflecting the degree of aversion (or proneness) to inequality in the distribution of health state utility between individuals (or groups) i and j. In the case of a

[9] Factors other than individual utilities are either regarded as irrelevant to social welfare or as being constant.

[10] Notice in this example that health state utilities have an upper bound of 1 (full health) and a lower bound of 0 (death). It would, however, be possible to expand this analysis to allow for states that attract negative utilities, i.e., for those considered to be worse than dead.

[11] In addition, it is possible that the objective is to equalise the health of both individuals (i.e., strict egalitarianism), which means the SWF would consist of points on the 45° line with points further from the origin being preferred to those closer to the origin. However, assuming that pathological budget lines or utility possibility frontiers (UPFs) are ruled out, preferences for strict egalitarianism will lead to the choice of the same point on the UPF as Rawlsian preferences.

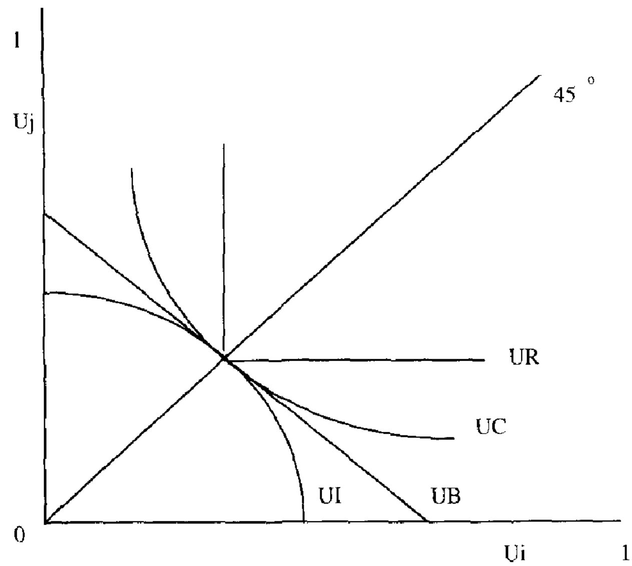

Fig. 1. Different formulatives of the social welfare function (SWF).

utilitarian HRSWF, i.e., UB in Fig. 1, A will take a value of 1 and for the Rawlsian case, UR, will be equal to negative infinity. Clearly, A will lie somewhere between these values for HRSWF (such as UC) that represent some trade-off between efficiency and equity. In the case of inequality proneness, as shown by UI in Fig. 1, A will be greater than 1. The parameter B determines the steepness of the iso-welfare loci, thereby reflecting the weight given to individual j relative to individual i. In Fig. 1, $B = 1$, which means that in every respect other than health, i and j are considered to be equal. [12]

4. Cobb-Douglas preferences

For illustrative purposes, let us assume that the HRSWF takes the log–linear form, such that $A = 0$. In such circumstances, the HRSWF is analogous to a Cobb–Douglas (CD) utility function

$$U(u_i, u_j) = u_i^{\alpha} u_j^{(1-\alpha)} \tag{6}$$

where a lies in the [0,1] interval. [13] Thus, in terms of the formulation in Eqs. (4) and (5), $B = (1 - a)/a$. CD preferences are the standard example of indifference

[12] In the original Atkinson formulation, it is assumed that each individual is treated symmetrically (by assuming two anonymous individuals whose (different) income levels are all that is known). In this way, B forms no part of the original formulation.

[13] Given that u_i and u_j are both between 0 and 1, $U(u_i, u_j) < 0$, but this can be changed by adding a constant, c (where $c > 1$), to u_i and u_j.

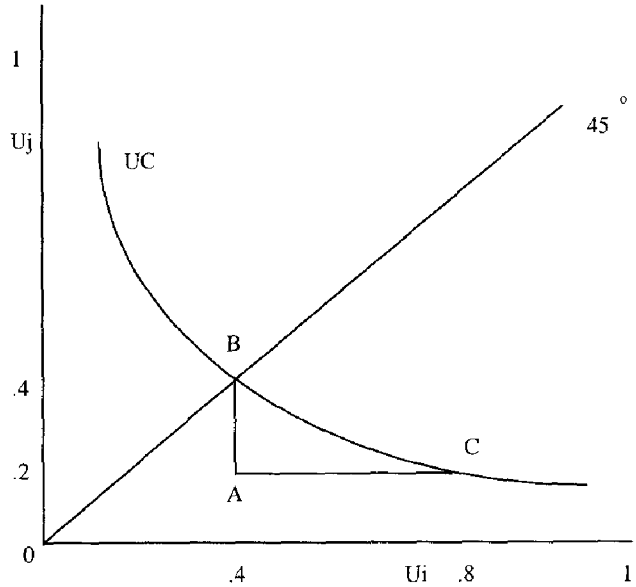

Fig. 2. Example of the Cobb–Douglas SWF.

curves that look well-behaved; in terms of a HRSWF they imply an aversion to inequality.

In Fig. 2, a is assumed to be 0.5 (i.e., $B = 1$), which means that the same weight is given to individual i as to individual j, perhaps because in every other respect (save health) they are considered to be equal. Consider an initial point such as A which results in a health state utility of 0.4 for individual i and 0.2 for j. Suppose that, given resource constraints, it is only possible to treat one person. We can either improve the health of j from 0.2 to 0.4 (i.e., move to point B) or improve the health of i from 0.4 to 0.8 (i.e., move to point C). An individual (or society) with CD preferences that give the same weight to each individual will be indifferent between these two alternatives. [14] Note that to be indifferent between the two alternatives, the health gain of the healthier individual, i, $(0.8 - 0.4 = 0.4)$ has to be greater than the health gain of the sicker individual, j, $(0.4 - 0.2 = 0.2)$, representing a distributional consideration.

The convenience of assuming CD preferences is highlighted when a (linear) budget constraint is introduced into the model. It is then straightforward to show that the optimal choices that satisfy this type of HRSWF are

$$u_i = \alpha\left(\frac{m}{p_i}\right), u_j = (1 - \alpha)\left(\frac{m}{p_j}\right) \tag{9}$$

[14] Given $a = 0.5$, $U(u_i, u_j) = 0.5 \ln u_i + 0.5 \ln u_j$. We already have one point, B, on the indifference curve that yields a utility of 0.5 ln (0.4) + 0.5 ln (0.4) and we know that to be on the same indifference curve, that this must be equal to 0.5 ln u_i + 0.5 ln (0.2). Rearranging and solving for i gives $i = 0.8$.

where m is the size of the total budget and p_i and p_j are the (constant) costs per unit of utility of treating i and j, respectively. Rearranging these formulae gives

$$u_i p_i = \alpha m, \; u_j p_j = (1-\alpha) m \tag{10}$$

This means that a fixed fraction of the health care budget is spent on each individual. The size of the fraction is determined by the exponent in the CD function (i.e., by the value of B in the class of HRSWF specified in Eqs. (4) and (5)).

Clearly, this property of CD preferences is useful when considering the distribution of health care expenditure between two individuals who are not considered to be equal. If we assume that individual j is given greater weight than individual i [15], then $a < 1 - a$ (i.e., $B > 1$) in the specification of the HRSWF. For example, if $a = 0.2$ (hence $1 - a = 0.8$) then a CD HRSWF takes the form $U(u_i,u_j) = 0.2 \ln u_i + 0.8 \ln u_j$. Thus, we know that society wishes to allocate 20% of its health care budget to the treatment of individual i and 80% to the treatment of individual j. Because social preferences (for fixed values of p_i and p_j) are a linear function of m, this will be true irrespective of the overall size of the health care budget.

We know from our earlier example (assuming $a = 0.5$) that a health gain of 0.2 to individual j was equal to a health gain of 0.4 to individual i. For the budget line, or the health-related utility possibility frontier (HRUPF), to pass through these two points requires the slope of the HRUPF to be -0.5, i.e., we can gain 0.5 units of health for j for every one unit of health we gain for i. [16] At this rate of transformation, a society with CD preferences would be indifferent between treating i and j. If the gradient of the HRUPF were to be steeper than -0.5 (for example, treating individual i becomes relatively more expensive), point B would be chosen, if it were flatter (for example, treating individual i becomes relatively cheaper), point C would be chosen. Assuming a continuous HRUPF [17] between B and C, we can maximise HRSW by doing something for both individual i and individual j. In our example, tangency between the HRSWF and the HRUPF is

[15] Previous research has shown that the general public may wish to weight more heavily the health needs of the young, those with children, and those who have looked after their own health (Charny et al., 1989).

[16] The strong assumption here is that there are no diminishing returns in the treatment of individuals i and j; in other words, the relationship between costs and health gain is linear. While diminishing returns is likely to be a more valid assumption (resulting in a UPF that is concave to the origin), for expositional purposes, constant returns are assumed.

[17] In health care, the UPF is most likely discrete rather than continuous, i.e., we can either improve the health of one person or we can improve the health of the other. But in some cases, a trade-off may exist.

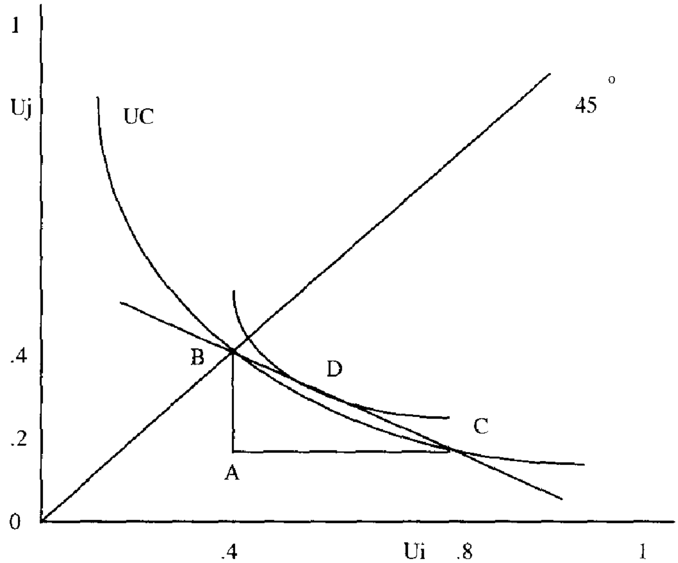

Fig. 3. Example of the Cobb–Douglas SWF with budget line.

where i's health improves from 0.4 to 0.6 and j's health improves from 0.2 to 0.3, as shown by point D in Fig. 3. [18]

5. Discussion

The class of HRSWF that has been postulated here is sufficiently comprehensive to encompass a wide range of prescriptions concerning the distribution of (health state) utility. In this paper, the HRSWF is characterised by two parameters; A, reflecting aversion (or proneness) to inequality in the distribution of utility, and, B, reflecting the relative weight given to the treatment of different individuals or groups.

Particular attention has been given to the log–linear form ($A = 0$) that implies inequality aversion and is analogous to a Cobb–Douglas utility function. This is largely for expositional purposes (CD preferences are well-behaved and have a number of useful properties) but partly because they enable us to represent various kinds of trade-off that society may be prepared to make between efficiency, in terms of health gain, and equity, in terms of severity of (pre- and post-treatment)

[18] We know that $u_i = 0.5\ \mathrm{m}/1$ and $u_j = 0.5\ \mathrm{m}/2$, so, if $u_i = 0.4$ and $u_j = 0.4$, we know that 1 is spent on i and 2 is spent on j, and if $u_i = 0.8$ and $u_j = 0.2$ we know that 1.6 is spent on i and 0.8 is spent on j. In other words, $m = 2.4$. For tangency with a higher SWF, we know with CD preferences that the budget is spent in proportion to the exponents a and $1 - a$, which when $a = 0.5$ means that 1.2 is spent on i and 1.2 on j. Substituting these values back into our original equations gives $u_i = 0.5 \times 1.2 = 0.6$ and $u_j = (0.5 \times 1.2)/2 = 0.3$.

illness and other relevant characteristics. In the absence of firm empirical evidence, it could be argued that a log–linear HRSWF (which, ceteris paribus, considers a health gain of 0.4 to an individual in a pre-treatment health state valued at 0.4 to be equivalent to a health gain of 0.2 to an individual in a pre-treatment health state valued at 0.2) is preferable to a utilitarian HRSWF that is concerned only with the maximisation of total health gain, irrespective of whether the person in the better or worse health state gets it.

If a fundamental prescriptive premise of conventional welfare economics is accepted, such that public sector decisions should reflect the strength of preferences of those who will be affected by those decisions, then it becomes an empirical question whether society is prepared to trade efficiency and equity against each other in this (or any other) way. In principle, this framework allows us to derive the precise shape of the HRSWF from responses to very simple questions.

Initially, the utility a respondent attaches to different states of health can be estimated using the SG or TTO. The x and y axes (i.e., the utilities of individuals i and j) can then be calibrated with health states that the respondent has valued for themselves. The respondent could be told that individuals x and y have preferences over health states that are identical to their own. They can then be asked questions along the lines of "if you had to choose between treating individual j who is in this health state (one the respondent valued at, say, 0.2) before treatment and this health state (one the respondent valued at, say, 0.4) after treatment or individual i who is in this health state (one they valued at, say, 0.4) before treatment and this health state (one the respondent valued at, say, 0.6) after treatment, which one would you choose to treat?" The final health state of individual i can then be made better or worse depending on whether the respondent chooses to treat j or i, respectively. In this way, the value of A can be estimated. The value of B could be estimated by stating that i and j differ according to a characteristic other than health; for example, age or sex. [19]

In addition, if it is assumed that the respondent is not inequality-prone such that their iso-welfare loci over the treatment of two individuals are not concave (i.e., that A is less than or equal to 1), then responses to these types of questions can also be used to test the validity of individual utilities. For example, if a respondent strictly prefers moving someone from a health state utility of 0.8 to full health to moving someone else from a health state utility of 0.1 to 0.4, then it is unlikely

[19] It is assumed in answering these questions that respondents give honest answers and they do not respond in a strategic way. While it is improbable that responses to SG or TTO questions will be biased by strategic considerations, it is possible that in response to questions regarding the treatment of other people, respondents may give answers that they consider to more socially acceptable, or 'virtuous': for example, by indicating that priority should be given to the worst-off individual. Virtuous response is recognised as a source of bias in the willingness-to-pay literature, but its impact in the health state measurement field is largely unknown and, by its very nature, difficult to test for.

that the health state valuations elicited from this respondent can be treated as having interval scale properties with respect to health.

Alternatively, with respondents that are familiar with the concept of health status measurement and particularly with the notion that health states may lie on a continuum from full health to (or beyond) dead, it might be possible to present them directly with health state valuations rather than using health states that have an implied value. This was an approach taken with a convenience sample of undergraduate students at the University of Newcastle. The question asked and the results therefrom are shown in Appendix A. Of course, the results from this experiment should in no way be considered definitive, but they do suggest that the approach adopted is a feasible one.

6. Conclusion

It has been suggested by a number of economists that decisions about how to allocate scarce health care resources should be informed by the cost-per-QALY of the different alternatives. The recommendation is that resources should be directed towards interventions that yield low-cost QALYs and away from interventions that yield high-cost QALYs, thus implying that the objective of the health care system should be to maximise the number of QALYs gained, i.e., to be efficient. In this respect, it has been argued that the QALY approach fails to consider distributional issues that are known to be important in the context of health care.

The person trade-off technique to valuing health benefits has been proposed as one way to address the trade-off that may exist between efficiency and equity (defined in terms of any conscious departure from QALY maximisation). While welcoming the distinction that proponents of this method make between QALYs as measures of production and as measures of social value, the problem with this method is that it does not provide a framework to separate out the various aspects of outcome that are known to be important, such as severity of illness (both pre- and post-treatment), health gain and the numbers of persons treated.

An alternative approach put forward in this paper is based upon first measuring individual utility (using relatively well validated elicitation procedures such as the standard gamble or the time trade-off) and then postulating the shape of the health-related social welfare function by estimating the parameters in a (very general) class of social welfare function based on that first proposed by Atkinson (1970). This second stage involves asking questions relating to the health gain of other people who differ in terms of their initial and final health status, and perhaps in terms of other relevant characteristics (such as their age or sex). The approach has a number of advantages: (1) it is well-grounded in economic theory; (2) it provides us with flexibility in determining the HRSWF; (3) it makes the assumptions and implications of the resultant function transparent; (4) it allows us to validate individual preferences; (5) and preliminary experimentation suggests that it should be relatively simple to examine empirically.

Acknowledgements

The motivation for this paper came from discussions with Alan Williams; I benefited enormously from his comments on earlier drafts. I am also indebted to Erik Nord and Michael Jones-Lee who provided detailed comments on the paper and stimulated fresh ideas. I would also like to thank Graham Loomes and Angela Robinson for their helpful comments and suggestions.

Appendix A. An experimental test of the methodology

As an illustrative example of how the model presented in this paper can be investigated empirically, 35 third-year economics undergraduates at the University of Newcastle were presented with choices concerning the treatment of two individuals along the lines discussed in the paper. The students had taken an option in Health Economics, and as such were familiar both with the concept of health status measurement and the techniques that can be used to elicit valuations. Therefore, it was decided that these respondents could be presented directly with health state valuations rather than first asking them to value a set of health states, and then using those health states in subsequent questions.

In each of four seminars, 8 or 9 students were asked to imagine that there are two individuals, i and j, who have preferences over health states identical to their own and are the same in all relevant respects except health. They were told that, at the moment, i is in a health state valued at 0.4 and j is in health state valued at 0.2. Respondents were then asked the following question: ''Imagine that there is a treatment available which could move i to a health state valued at 0.6 or move j to a health state valued at 0.4. If you could only treat i or j but not both, who would you choose to treat?''

In total, one respondent preferred to treat i, stating that she felt it was better to have one person in a 'good' health state and one person in a 'bad' state rather than to have both in 'moderate' states. Two respondents were indifferent between treating i and j, stating that the health gain was the same in both cases. The remaining 32 respondents said they would prefer to treat j because it is 'fairer'; either because j is initially in a worse health state or because the distribution of health after treating j is more equitable than after treating i.

The 32 respondents who preferred to treat j were asked a second question: ''Imagine that, as before, the treatment will move j from a health state valued at 0.2 to one valued at 0.4, but that the treatment will now move i from a health state valued at 0.4 to full health (i.e., valued at 1.0). If you again had to choose between treating i and j, who would you choose to treat?''

Eight respondents still chose to treat j for both the reasons of 'fairness' cited above. Twenty-four respondents now chose to treat i on the grounds that the benefit to i is now much larger than the benefit to j. The 24 respondents who now chose to treat i were asked one final question: "Imagine again that the treatment will move j from a health state valued at 0.2 to one valued at 0.4 and that the treatment will move i from a health state valued at 0.4 to one valued somewhere between 0.6 and 1.0. Where between 0.6 and 1.0 would the treatment have to move i to, so that you are indifferent between treating i and j?"

The responses were as follows: 0.65 = 3; 0.70 = 7; 0.75 = 4; 0.80 = 5; 0.85 = 1; 0.90 = 2; 0.95 = 2. The median value for the full group of 35 respondents is 0.80 (inter-quartile range = 0.70–0.95). In other words, moving one person from a health state valued at 0.2 to one valued at 0.4, on average, yields the same social value as moving another person (who is identical in all respects except health) from a health state valued at 0.4 to one valued at 0.8. For this group of respondents, then, the log-linear HRSWF described in the paper would be a good approximation of their preferences.

References

Arrow, K.J., 1951. Social Choice and Individual Values. Wiley, New York.

Atkinson, A.B., 1970. On the measurement of inequality. J. Economic Theory 2, 244–263.

Bergson, A., 1938. A reformulation of certain aspects of welfare economics. Q. J. Economics 52, 310–334.

Charny, M.C., Lewis, P.A., Farrow, S.C., 1989. Choosing who shall not be treated in the NHS. Social Sci. and Medicine 28 (12), 1331–1338.

Culyer, A.J., 1995. Equality of what in health policy? Conflicts between the contenders. Centre for Health Economics Discussion Paper No 142. Univ. of York.

Dyer, J.S., Sarin, R.K., 1982. Relative risk aversion. Management Sci. 28 (8), 857–886.

Gravelle, H., Rees, R., 1981. Microeconomics, Longman, London.

Jones-Lee, M.W., Loomes, G.C., 1995. Discounting and safety. Oxford Economic Papers 47, 501–512.

Mooney, G., Olsen, J.A., 1991. QALYs: where next? In: McGuire A., Fenn, P., Mayhew, K. (Eds.), Providing Health Care: the Economics of Alternative Systems of Finance and Delivery. Oxford Univ. Press.

Ng, Y., 1983. Welfare Economics: Introduction and Development of Basic Concepts. MacMillan.

Nord, E., 1994. The QALY: A measure of social value rather than individual utility. Health Economics 3 (2), 89–93.

Nord, E., Richardson, J., Macarounos-Kirchmann, K., 1993. Social evaluation of health care versus personal evaluation of health states: evidence on the validity of four health state scaling instruments using Norwegian and Australian surveys. Int. J. Technol. Assessment 9 (4), 463–478.

Pareto, V., 1935. The mind and society, vol 4, edited by Livingston A.

Rawls, J., 1971. A Theory of Justice. Harvard Univ. Press.

Sen, A.K., 1977. On weights and measures. Econometrica 44, 1539–1572.

Sugden, R., 1981. The Political Economy of Public Choice. Martin Robertson, Oxford.

Sugden, R., Williams, A., 1978. The Principles of Cost-benefit Analysis. Oxford Univ. Press.

Wagstaff, A., 1991. QALYs and the equity–efficiency trade-off. J. Health Economics 10, 21–41.

Weinstein, M.C., Stason, W.B., 1977. Foundations of cost-effectiveness analysis for health analysis and medical practices. New England J. Medicine 296, 716–721.