

The Significance of Distributive Effects in Social Assessment of Health Care*

Eva Rodríguez[†] and José Luis Pinto

University of Vigo

Pompeu Fabra University

*The authors would like to thank Han Bleichrodt for his comments and suggestions. Any remaining errors are the authors' responsibility. José Luis Pinto acknowledges financial support from the Dirección General de Enseñanza Superior, SEC 98-0296-C04-01. Eva Rodríguez acknowledges financial support from the University Complutense of Madrid in the form of grant received while this study was conducted.

[†]Correspondence to: Departamento de Economía Aplicada, Facultad de Economía, Universidad de Vigo, Lagoas-Marcosende s/n, 36200 Vigo, Galicia, España. Tel: + 34 986 812 516; fax: + 34 986 812 401; E-mail: emiguez@uvigo.es

Abstract

In this paper we address the importance of distributive effects in the social valuation of QALY's. We propose a social welfare function that generalises the functions traditionally used in the health economic literature. The novelty is that, depending on the individual health gains, our function can represent either preferences for concentrating or preferences for spreading total gain or both together, an issue which has not been addressed until now. Based on an experiment, we observe that this generalisation provides a suitable approximation to the sampled social preferences.

KEY WORDS: QALY; distributive effects; Social Welfare Function; inequality aversion.

JEL: I18, D61,D63

1 Introduction

Cost-effectiveness analysis (CEA) is a methodology which aims to that facilitate social decision-making in the allocation of scarce resources. When applied in the health care sector, it requires an output measure which compares the effects of different health care programs on people's health. The QALY —quality-adjusted life year— has been proposed as an adequate outcome measure since it combines quantity and quality of life in one single index. However, given its theoretical foundations in Utility Theory,^{[1],[2]} the QALY is considered by some authors to be a utility measure. So, CEA is sometimes called cost-utility analysis when effects are assessed using QALYs.

By calculating cost per QALY, it is possible to compare different health care programs in terms of their efficiency, and it has been argued that the best allocation of health resources is that which maximises the community's health, as measured by the unweighted sum of individual QALYs. In this way, each additional QALY is implicitly considered to have the same social value, independent of the characteristics of the patient and the number of QALYs received. We will refer to this unweighted sum model as the aggregated QALY model (AQM).

Given that a QALY is always assigned the same social value, giving many

QALYs to a few people will have the same social value as giving a few QALYs to many people, as long as the total number of QALYs provided remains constant. However, the AQM's failure to take into account distributive effects has incited great controversy. Some empirical studies^{[3],[4]} have shown that when the general public is asked to allocate health resources, they not only consider the total health gains for a given cost —efficiency— but also the way in which health gains are distributed among the population. Similar results have been obtained using the veil of ignorance approach.^{[5],[6]} Therefore, both aspects must be considered in the social assessment of health care programs.

Two principal means for incorporating distributional concerns into the QALY approach have been suggested. One has its roots in modern welfare economics^[7] and the other in multi-attribute utility theory.^[8] Both approaches propose variations in the aggregated QALY model which allow us to incorporate preferences for distributive effects. Bleichrodt^[8] suggests weakening the additivity condition underlying the model and formulates, under uncertainty, a multiplicative model based on individual QALY gains. Starting from the parameters of the multiplicative function, it is possible to analyse the extent to which society is willing to sacrifice efficiency in order to obtain a more equitable distribution.

Wagstaff,^[7] while not giving up the additivity condition, proposes an isoelastic social welfare function inspired in the function proposed by Atkinson.^[9] In this case, distributive preferences are introduced by designating decreasing social values to each additional QALY received by the same individual. Dolan^[4] also analyses the properties of this function and its application in the allocation of health resources. Another way to introduce this decreasing social value in an additive function is by using a “social weight rate” as proposed by Olsen.^[3] Both additive propositions will be analysed in greater detail in the third section. We use the additive approach here because it allows us to use concepts which are commonly employed in the inequality literature, and because it will allow us to compare the results of our research with previous empirical findings.

Parallel to the theoretical debate, some empirical studies have attempted to estimate parameters for the social welfare function (SWF) which provide the best fit with sampled social preferences. For example, Johannesson and Gerdtham^[5] estimated the shape of the SWF using the veil of ignorance approach, and Olsen^[3] estimated the social weight rate under certainty. However, these studies did not consider the possibility that the parameters might vary as a function of the number of QALYs gained, or took this into account

in only a limited fashion, by comparing few and similar gains.

However, there might be a stronger preference for a more equitable distribution when individual gains are very great than when they are small. In this case, the parameters will not be constant. Furthermore, there is some empirical evidence which suggests that people may prefer to concentrate gains when only a small number of individual QALYs are obtained. For example, Pinto and Lopez^[10] showed that people preferred concentrate gains when comparing small quality of life increments obtained with life saving treatments. Choudry and colleagues^[11] also reported that people preferred a program that increased life-expectancy by 20 years in 500 people over a program that increased life-expectancy by 1 year in 10,000 people. In other words, there may be a preference for substantial improvements in a few individuals over “insignificant” improvements for many. Therefore, a more flexible functional form of the SWF may be needed which, allows us to describe possible changes in the pattern of preferences. This is the primary aim of this study.

In following section, we identify some conditions, frequently used in welfare economics under certainty, which are compatible with the AQM and with more flexible models. In section 3, we propose two measures which will allow

us to calculate the degree of inequality aversion. Based on these inequality aversion measures, the differences between the two previously mentioned additive SWFs, will be analysed. We finish the section by proposing a specific SWF that generalises the SWF underlying the AQM and lets us introduce different distributive preferences. Section 4 shows the results of an experiment we carried out to construct a SWF that best fits the preferences of the respondents and to analyse whether the degree of inequality aversion is independent of the gains provided or not. Section 5 discusses the results obtained. Finally, section 6 contains concluding remarks.

2 Derivation of the Social Welfare Function

The aggregate QALY model, and variations of it which have attempted to introduce distributive considerations, can be derived from an additive SWF where social welfare is defined as a function of individual health gains. In order to generate this SWF, some conditions must first be established—for a more detailed and formal exposition of these conditions, see Rodríguez and Pinto.^[12]

The output of a given health care program is defined as a distribution of

health gains, measured as the number of QALYs the program provides to a given population. Let $n \geq 3$ be the population size and let $T \in \mathfrak{R}_+^n$ be the set of possible outputs resulting from the implementation of different health care programs. An element of T is defined as a vector, $\tau = (t_1, \dots, t_n)$, where $t_i \in \mathfrak{R}_+$ — $i = 1, \dots, n$ — indicates the number of QALYs individual i receives from the program.

The next step is to establish a criterion of social choice which allows us to order all the elements of T unambiguously. In order to do this, we consider that the social preference relationship is complete and transitive. Thus, it can be represented by a value function —SWF— defined over τ , that we denote by $W(\tau)$. In addition, the SWF is considered to depend positively on individual QALYs — $W(\cdot)$ increases in t_i (Pareto assumption).

It would seem appropriate to assume that the social preferences for two different distributions of health, which differ only in the amount of health gain for two random individuals, depends only on the number of QALYs received by those two individuals (independence assumption).

These assumptions allow us to define an additive SWF in the following way $W(\tau) = \sum_{i=1}^n u_i(t_i)$, where u_i is a positive monotonic transformation defined over t_i , that reflects the interpersonal comparisons made by

society.^{[13],[14]}

Another frequently used assumption is that of anonymity.^[8] This assumption tells us that if a health distribution is a permutation of another distribution, then both distributions must have the same social value. Based on this assumption and a scaling assumption, we can define the SWF as,

$$W(\tau) = \sum_{i=1}^n u(t_i) \quad (1)$$

where $u(t_i)$ indicates the social utility of the t_i gain.

Function $W(\tau)$ is compatible with different social preferences depending on the functional form of $u(t_i)$. Suppose that $u(t_i)$ is a continuous and twice differentiable function. Given that $u'(t_i)$ defined as $du(t_i)/dt_i$ represents (social) marginal utility, it can be interpreted as the social weight designated to each additional unit of t received by individual i .^[15] The reason for the latter affirmation is as follows. If a health care program produces a small change in everyone's health, $(\Delta t_1, \dots, \Delta t_n)$, social welfare will rise, $\Delta W = \sum_{i=1}^n u'(t_i)\Delta t_i$. Therefore $u'(t_i)$ acts as a system of weights when summing the effects of the program over the whole population.

The SWF that underlies the AQM is reached immediately if we impose a restriction on the marginal utilities. Given that, for this model, an additional

QALY always has the same social value, $u'(t_i)$ must be constant. Under this assumption, $u^{(1)}(t_i) = t_i$ —or any positive linear transformation— and the SWF —Eq.(1)— can be defined as,

$$W^{(1)}(\tau) = \sum_{i=1}^n t_i. \quad (2)$$

In order to introduce the existence of a temporal discount rate we can simply suppose that t_i are QALYs that have already been discounted.

3 Distributive considerations

What happens if people who have to choose between different health programs are concerned not only with the total QALYs provided, but also with how those QALYs are distributed? In this case, each additional QALY received by the same individual may have different weights attached, so that $u'(t_i)$ varies depending on the value of t_i . If we suppose that society prefers a more equitable distributions —positive inequality aversion—, the weight of each additional QALY received by individual i , will decrease as the number of QALYs received increases. In this case, $u''(t_i)$ defined as $du'(t_i)/dt_i$ is negative — $u(t_i)$ is a concave function. If, on the other hand, there are pref-

erences to concentrate gains —negative inequality aversion—, then $u''(t_i)$ is positive or, equivalently, $u(t_i)$ is a convex function.

Once we have defined the sign of aversion, we address the question of how to measure the degree of inequality aversion. For this is important to have some measure which allows us to determine the extent to which society is willing to give up a certain degree of efficiency for alternative distribution of health gains. There are two measures which are particularly appropriate, given the cardinality of $u(t_i)$ —the expression $u''(t_i)$ is not adequate because it is not invariant to positive linear transformation of $u(t_i)$. One is a measure of absolute inequality aversion, $\theta_a(t_i) = -u''(t_i)/u'(t_i)$, and the other is a measure of relative inequality aversion, $\theta_r(t_i) = -t_i [u''(t_i)/u'(t_i)]$. It should be noted that the latter measure is the elasticity of marginal utility and indicates the percentage reduction in the weight of each person, $u'(t_i)$ when the number of QALYs is increased by 1 percent. Both measures have their origins in Arrow^[16] and Pratt's^[17] aversion to risk measures.

Given that $u'(t_i)$ is always positive, both measures are positive if society prefers to distribute health gains, they are negative if there are preferences for concentration, and they are equal to zero if only the total gain matters (AQM). On the other hand, a constant θ_a indicates that in the presence of

equal changes in patient health level—in this case number of QALYs—, the weight is modified by the same proportion, independently of the value of t_i . However, a constant θ_r reveals that the weight is modified in the same proportion in the presence of equal proportional changes.

Inequality aversion measures allow us to analyse those assumptions that underlie different SWFs. As mentioned in the introduction, in the literature on QALYs, under certainty, two ways are normally used to include distributive preferences starting from an additive SWF. Wagstaff proposes an isoelastic SWF that, along with the anonymity assumption, defines $u(t_i)$ as

$$u^{(2)}(t_i) = \begin{cases} (1 - \varepsilon)^{-1} t_i^{(1-\varepsilon)} & \text{if } \varepsilon \neq 1 \\ \ln t_i & \text{if } \varepsilon = 1 \end{cases} \quad (3)$$

Olsen proposes using a social weight rate, $1/(1+r)$, when it comes to assessing each additional QALY. In this case, $u^{(3)}(t_i) = \sum_{j=1}^{t_i} [1/(1+r)]^j$. We assume that discounted QALYs are used, otherwise the rate $1/(1+r)$ reflects the social weight rate and the temporal discount rate in an indistinguishable way. Given that we have considered that $t_i \in \mathfrak{R}_+$, the continuous version of $u^{(3)}$ is expressed as follows,

$$u^{(3)}(t_i) = \int_0^{t_i} e^{-rt} dt = \frac{1}{r} (1 - e^{-rt_i})_i. \quad (4)$$

It is clear that positive (negative) ε and r values correspond to concave (convex) utility functions and therefore describe a positive (negative) inequality aversion. If ε and r are equal to zero it would indicate that maximising the number of QALYs is the only consideration of interest, therefore, $u^{(2)} = u^{(3)} = u^{(1)}$.

The main difference between the two functions is determined by their degree of inequality aversion. While the weights designated by $u^{(2)}$ increase in the same proportion in the presence of identical proportional changes — decreasing θ_a and constant θ_r — the weights designated by $u^{(3)}$ increase in the same proportion in the presence of equal changes in the level —constant θ_a and increasing θ_r .

Given that recommendations regarding health policy may be very different if one or another function is used, it is necessary to know which best reflects social preferences. Furthermore, it ought to be borne in mind that if there are different degrees of both absolute and relative inequality aversion, depending on the value of t , then neither of the two previously mentioned functions would be valid. In this case it would be necessary to apply more flexible functions.

In this context, it would be appropriate to define a social utility function,

denoted as $u^{(4)}(t_i)$, that generalises the above mentioned formulations so that they can be obtained as a particular case of $u^{(4)}$. A function that fulfils these requirements can be defined as,

$$u^{(4)}(t_i) = \alpha_1 e^{-\alpha_2 t_i} t_i^{\alpha_3}. \quad (5)$$

This function can have concave or convex sections depending on the value of the parameters. This is an important property because it allows us to represent social preferences with positive and negative inequality aversion in the same function and therefore permits us to represent a change in the preference pattern. In addition, if $\{\alpha_1, \alpha_3\} = 1$ and $\alpha_2 = 0$ then $u^{(4)} = u^{(1)}$; if $\alpha_1 = 1/\alpha_3$ and $\alpha_2 = 0$ then $u^{(4)} = u^{(2)}$; finally, if $\alpha_1 = -1/\alpha_2$ and $\alpha_3 = 0$, $u^{(4)}$ will be a linear transformation of $u^{(3)}$.

4 Experiment

The aim of this experiment was to obtain a first approximation of the function $u(t_i)$ and, therefore, of the SWF —Eq.(1). In order to do this, a set of health gains which we consider representative ex ante are assessed. Then the functional form that best fits these assessments is sought. Once $u(t_i)$ is

obtained, its properties are analysed, with the focus being on the influence that distributive effects have on the assessment of any health gain.

The Person Trade-Off (PTO) technique^[18] was used to assign social values to individual health gains. Briefly, this technique consists of presenting respondents with different allocations of numbers of patients and the health gains they receive to determine which allocations are equally preferred by the respondent.

To avoid unnecessary notation sub-index i will henceforth be omitted.

4.1 Design

The experiment was conducted on 61 undergraduate students —21 Economics students, 20 Political Science students and 20 Law students. The students were paid approximately \$16 for their participation. The experiment consisted of three meetings with the participants on three different days. At the first meeting the aim of the study was explained to the participants. They then filled out a pilot questionnaire to familiarise them with the kind of questions they would be asked at the second meeting.

The second meeting was carried out in different sessions with an average of five participants per session. Each individual was shown different health care

programs that were all directed at 20-year-old patients. By using the same age group we tried to avoid any potential effect of patient age on decision-making. Each program consisted of a different pair of values (t, p) , where p is the number of patients who would benefit from its application, and t is the health gain, measured in life years in full health, that each patient would receive. Since the number of life years are in full health they can be interpreted as the number of QALYs. For each program, the participant had to say how many 20-year-old patients, p^* , would have to receive a 10-year increase in life-time to make him indifferent between both programs. In other words, once t and p have been fixed, they must give p^* a value such that he or she is indifferent between (t, p) and $(10, p^*)$, i. e. the PTO technique. Given that in the pilot study participants were found to have some difficulties in choosing a concrete number of years, “choice-bracketing” was used to calculate the p^* value. This mechanism consists of approaching the value through a series of successive questions where choices have to be made between two allocations —see appendix. When it was not possible to obtain an exact value for p^* using the choice-bracketing technique but an interval, the intermediate value of the interval was used.

Our working hypothesis was that preferences for distributive effects could

vary depending on the amount of life years received by each patient. To test this, participants assessed five different life–time increases: 1, 2, 5, 20 and 50 years. The number of patients, p , was selected in such a way that all of the programs provided the same total increase in life years and therefore had the same value within the AQM. Accordingly, they assessed the following programs: $(1, 100)$, $(2, 50)$, $(5, 20)$, $(20, 5)$, $(50, 2)$, where the first element refers to the increase in life years for each individual and the second element refers to the number of patients who receiving that increase.

In any experiment of this kind it is important to analyse the extent to which using another technique provides similar results, to determine consistency across methods. Therefore, after applying the choice–bracketing technique to all allocations, we provided each participant with six cards that they had to rank from more to less preferred (the direct ordering technique (DO)), each card corresponding to one of the previously assessed programs. The additional card corresponded to that program in which life was increased by ten years for ten patients — $(10, 10)$. It did not make any sense to assess this allocation before because the ten–year increase was used as the reference program against which the others were compared. Finally, participants were asked to briefly justify their ranking.

Two weeks later, we organised a third meeting in which the experiment was repeated to check whether the results were consistent over time (test-retest reliability).

4.2 Method of analysis

Individuals who did not make trade-offs were excluded from the analysis. Next, from valuations of participants obtained with the PTO technique, we calculated the average value of p^* for each (t, p) pair and determined whether each p^* mean value was statistically equal to ten using Student t test. If the AQM assumptions appropriately describe preferences for resource distribution then p^* mean values should be equal to 10.

Based on individual p^* , we obtained the social value assigned by each participant to the five increases in question. If we assume that the social value of increases in life years is independent of population size^[3] we get $u(t) * p = u(10) * p^*$. Given that the utility function can be scaled arbitrarily we get $u(10) = 10$, so that the social value of each time increase, t , is expressed as

$$u(t) = 10 * p^*/p. \tag{6}$$

Lastly, we look for the functional form, $\hat{u}(t)$, that best fits with these

values, the aim being to obtain assessments for gains not assessed directly. To avoid imposing restrictions on the model, different regressions were estimated using Ordinary Least Squares to find that with the best goodness of fit, as measured by the adjusted (for degrees of freedom) R^2 .

To test the correlation between rankings obtained using direct ordering and PTO at the individual level, a Spearman rank correlation coefficient (SCC) was calculated for each participant, and the mean SCC was calculated for all participants.

To analyse correlation at the social level, the individual rankings obtained using both techniques were aggregated using the Borda rule. In this way, two social orderings were obtained, denoted as S-PTO and S-DO. To assess the degree of correlation between both orderings, we applied both SCC and Kendall rank correlation coefficient (KCC). The KCC is not used to evaluate the correlation at individual level because it is necessary that any card has the same position in the ranking. However, we observed that this did not occur at the individual level.

To analyse the correlation between the ranking from the initial DO and that from the retest, we use SCC. To analyse the correlation between valuations resulting from the initial PTO technique and the retest, we used

Pearson linear correlation coefficient (PCC).

4.3 Results

Sixteen of the 61 participants (26%) who did not make trade-offs were excluded from the analysis. Those excluded always chose the pairs with the greatest number of patients (10 participants) or the pairs with the greater number of years (6 participants).

Table 1 shows the average social value for each allocation of life years increases and number of patients. As can be seen, the hypothesis $p^* = 10$ is rejected at the 1% (three cases), 5% (one case) and 10% (one case) level. Thus Eq.(2) does not accurately represent the distribution preferences found here.

Using the mean values it is possible to analyse the distributive preferences. We must bear in mind that given two allocations (t', p') and (t'', p'') , whose p^* values are $p^{*'}$ and $p^{*''}$, respectively, the participants prefer to distribute (concentrate) gains if, when $p' > p''$ then $p^{*'} > p^{*''}$ ($p^{*'} < p^{*''}$). Table 1 shows that they prefer to distribute gains in some cases, for instance when comparing 20 and 50 years, but that they prefer to concentrate gains in other cases, for instance in the case of 10 and 2 year increases.

The individual p^* 's provided the social value that each participant assigned to each time increase —Eq.(6). Using these values, it was possible to determine which function best reflected participant preferences; in this case the function with the best goodness of fit was the $u^{(4)}$ function —Eq.(5). Assuming a multiplicative error term in this last equation, we can linearize by means of a log transformation, obtaining the following results:

$$\widehat{\ln u(t)} = - \underset{(-7.19)}{0.807} - \underset{(-6.55)}{0.026} t + \underset{(21.08)}{1.435} \ln t, \quad R^2 (adjusted) = 0.85, \quad (7)$$

The $t - ratio$ in brackets was calculated using a robust heteroscedastic covariance matrix estimator.

Rewriting Eq.(7) in its original form, we get

$$\hat{u}(t) = 0.446 e^{-0.026 t} t^{1.435}. \quad (8)$$

The functional form of $\hat{u}(t)$ allows us to analyse distributive preferences. Starting from a simple derivation exercise, we find that $\hat{u}''(t)$ is positive for t values under 9.1 and negative for the remaining feasible values. Figure 1 represents the function $\hat{u}(t)$, which starts out being slightly convex but becomes concave when the t value is 9.1. This information has important qualitative implications. When gains are under 9.1 years, participants on average prefer to concentrate those gains, but if the gains are greater than

this *threshold* gain they prefer to distribute them.

Using Eq.(8) and Eq.(1) we obtain the associated SWF value for health care program $\tau = (t_1, \dots, t_n)$,

$$\hat{W}(\tau) = \sum_{i=1}^n \hat{u}(t_i) = 0.446 \sum_{i=1}^n e^{-0.026 t_i} t_i^{1.435}. \quad (9)$$

As Wagstaff and Dolan report, the indifference curves of the SWF provide another interesting way to analyse distributive preferences. Figure 2 shows the indifference curves of $\hat{W}(\tau)$ supposing $\tau = (t_1, t_2)$. For individual values under 9.1 (area I) indifference curves were concave meaning that the overall preference is to concentrate gains: for any given amount of health gains, allocations that concentrate gains are always placed on a higher indifference curve than those which distribute them. For larger increases (area II) convex indifference curves indicating preferences for more equitable distribution can be seen. We cannot say anything a priori about (symmetrical) areas III and IV. Given that both of these areas combine individual values where there is a preference for spreading gains with individual values where there is a preference for concentrating gains, the final result will depend on the specific gains in question.

Following the theoretical exposition, the measures of absolute and rela-

tive inequality aversion — θ_a and θ_r respectively— were calculated. Obviously, both measures are negative for values lower than 9.1, indicating the existence of negative inequality aversion. Above this value, both parameters are positive thereby indicating positive inequality aversion. It is interesting to analyse the trajectory of both indicators. It is easy to verify that $d\theta_a(t_i)/dt_i$ and $d\theta_r(t_i)/dt_i$ are positive for all t_i analysed. Given that both coefficients are increasing with respect to health gain, the greater the number of years provided to each individual, the greater the inequality aversion, in both absolute and relative terms.

Table 2 shows the S-PTO, the S-DO, and the SCC and KCC between both rankings. In addition, the average SCC of all participants is shown. Both methods provide similar orderings, and therefore high correlation coefficients, suggesting substantial consistency across methods both at the social and individual level.

Finally, test-retest reliability for direct ordering was 0.93 as measured using Spearman's correlation coefficient, and 0.44 for the PTO using Pearson's correlation coefficient. This suggests a high stability of preferences with regard to the ranking, but a weaker stability with regard to the actual values obtained with the PTO.

5 Discussion

The results of the experiment performed here suggest that distributive preferences can vary in intensity and may even change direction depending on the individual health gain provided. Participants, on average, prefer health programs which distribute benefits over a greater number of people, provided that the gain to each patient is *sufficiently* high. In this experiment, a gain was considered sufficiently high when it was over 9 years —the threshold gain. For example, in this experiment programs that provided a gain of 15 years to two patients, were more highly valued than programs providing 10 years to one patient and 20 years to the other. These kind of results reflect social preferences for spreading gains over many people, thus incorporating a higher degree of social justice. This inequality aversion is more intense the higher the individual health gain.

On the other hand, respondents preferred to concentrate gains in a smaller number of individuals, when gains to each individual were below the threshold gain of 9 years. For example, providing one year life time increases to 8 patients is less highly valued than providing 8 years of additional life to one patient. A possible explanation for the existence of this threshold gain is that participants consider shorter periods of time to be less worthwhile in

the sense that they would not warrant initiating new life projects. Moreover, as Choudhry and colleagues^[11] point out, “large gains are more visible even if restricted to fewer people.”

The proposed SWF allows us to represent this type of preference and suggests a need to reconsider the underlying assumptions of CEA with regard to the valuation of health output. It may be, for instance, that some health programs which are rejected because of their low cost–effectiveness ratios, would be approved if distributional preferences were taken into account. For example, Schapira and colleagues^[19] estimated the effectiveness of a specific Ovarian Cancer Screening and found that for every 100,000 women screened, on average 14 would test positive and that their lives would be prolonged 11.6 years each. Given that the number of days of extended life provided, on average, to each women screened is 14 hours, they concluded that “mass screening for ovarian cancer will not improve average life expectancy in the population by a meaningful amount of time and cannot be recommended as an effective health policy.” For the authors —and in CEA in general— 14 hours to 100,000 people is equivalent to 11.6 years each to 14 people. However, an increasing number of critics have questioned this assumption and have argued that the second combination should be given a higher social

value. Our SWF does in fact do this, with 14 hours each to 100,000 women receiving a social value, W , of 4.3 and 11.6 years each to 14 women having a value of 155.6.

Our results may also have more general implications, although these should be tested empirically. For example, our results may help to explain why programs that considerably improve the quality and/or the quantity of life of a small number of patients (e.g. organ transplantation), may be preferred to programs which give very small gains to too many patients (e.g. dental fillings), even when the former have a lower cost effectiveness ratio.

Finally, it is important to stress the pilot nature of this experiment and therefore its limitations. In particular, the sample used was a convenience sample, and the experiment should be repeated in a more representative sample to test the robustness of the results. Second, to avoid confusion between distributional preferences and age weights all patients in the experiment were 20 years old. However, it is possible that distributional preferences will differ depending on the patient's age. Thus, for example, the threshold gain could be smaller for older participants. Third, the estimated $u(t)$ starts to decrease when the gain received by 20-year old a patient is (for life-time increases) over 55 years. This is inconsistent with the Pareto assumption, and with

common sense. However, this anomaly is not a limitation of the function proposed but of the experiment. It should be noted that the biggest gain in life years evaluated was 40 years, so our results are only relevant for a gain of up to this size. Finally, to estimate $u(t_i)$, a number of participants were excluded because they refused to make trade-offs, and the resulting function does not take their preferences into account.

6 Conclusion

It has been suggested by a number of economists that distributive preferences should be included in the social valuation of QALYs. In this paper, we have proposed a SWF that allows us to represent different social preferences, including the additive ones proposed in the literature on QALYs to date. Moreover, this function allows us to combine preferences for concentration and distribution, an issue explored by some authors, but which had not been formulated theoretically.

Our results confirmed the hypothesis that distributional preferences depend on the size of health gain. Participants preferred programs which distributed the total gain as long as they provided a *sufficiently* big individual

gain, but they preferred to concentrate the gain rather than give *insignificant* gains to many people. The SWF proposed in this paper provides a suitable approximation to the preferences of the group studied, given that it allows us to combine concave and convex sections of the utility function.

Although this paper throws light upon the way that distributive effects can be introduced into the social valuation of health output, more research and social debate is needed before these results can be used as a guide to health care policy making.

Appendix

Part of the questionnaire we used is shown below. One of the 5 time increase plus people allocations that participants assessed using choice–bracketing is included as an example.

In this section 2 treatments are described: A and B. The treatments differ from each other in the number of additional healthy life years provided to the patient, and in the number of people who receive gains. All patients are 20 years old. You must say whether you prefer treatment A, treatment B, or whether you are indifferent between them. Depending on your choice the questionnaire continues in the following way:

-If you choose an option where you find the word “stop”, circle the word and go on to the next table (in which treatment A has been varied).

– If you choose an option where you find the word “continue”, go on to the next line.

By way of simplification we will use the following notation:

Healthy life – year increases for the patient = “Years”

Number of people receiving gains = “People”

I prefer treatment A = “Pref. A”

I am indifferent to A and B = “Same”

I prefer treatment B = “Pref. B”

The treatments are as follows:

Treatment A

Treatment B

<i>Years</i>	<i>People</i>	<i>Years</i>	<i>People</i>	<i>Pref. A</i>	<i>Same</i>	<i>Pref. B</i>
5	20	10	1	continue	stop	stop
5	20	10	20	stop	stop	continue
5	20	10	3	continue	stop	stop
5	20	10	18	stop	stop	continue
5	20	10	5	continue	stop	stop
5	20	10	15	stop	stop	continue
5	20	10	8	continue	stop	stop
5	20	10	12	stop	stop	continue
5	20	10	10	stop	stop	stop

References

- [1] Torrance, G.W. Toward a utility theory foundation for health status index models. *Health Services Research* 1976; winter, 349–369.
- [2] PlisKin, J.S., Shepard, D.S. and Weintein, M.C. Utility functions for life years and health status. *Operations research* 1980, **28**: 206–24.
- [3] Olsen, J.A. Persons vs years: two Ways of eliciting implicit Weights. *Health Economics* 1994; **3**: 39–46.
- [4] Dolan, P. The measurement of individual utility and social Welfare. *Journal of Health Economics* 1998; **17**: 39–52.
- [5] Johannesson, M. and Gerdtham U–G. A note on the estimation of the equity–efficiency trade–off for QALYs. *Journal of Health Economics* 1996; **15**: 359–368.
- [6] Nord, E., Richardson, J., Street, A., Kuhse, H. and Singer, P. The significance of age and duration of effect in social evaluation of health care. *Health Care Analysis* 1996; **4**: 103–111.
- [7] Wagstaff, A. QALYs and the equity–efficiency trade–off. *Journal of Health Economics* 1991; **10**: 21–41.

- [8] Bleichrodt, H. Health utility indices and equity considerations. *Journal of Health Economics* 1997; **16** (1): 65–91.
- [9] Atkinson, A.B. On the measurement of inequality. *Journal of Economic Theory* 1970; **2**: 244–263.
- [10] Pinto J.L. and Lopez, A. More evidence on the plateau effect. *Medical Decision Making* 1998; **18**: 287–294.
- [11] Choudry, N., Slaughter, P., Sykora, K. and Naylor, C. Distributional dilemmas in health policy: large benefits for a few or smaller benefits for many. *Journal of Health Services Research and Policy* 1997; **2** (4): 212–216.
- [12] Rodríguez, E., Pinto, J.L. Social value of health programs: Is age a relevant factor?. *Health Economics* (forthcoming).
- [13] Debreu, G. Topological methods in cardinal utility theory. In: Arrow, K. J., Karlin, S. and Suppes, P. , eds. *Mathematical methods in the social sciences* 1960; (Stanford University Press) 16–26.
- [14] Keeney, R.L. and Raiffa, H. *Decisions with multiple objectives* 1976; (Wiley, New York).

- [15] Cowell, F.A. *Measuring Inequality* 1995; (Prentice Hall/Harvester Wheatsheaf).
- [16] Arrow, K. *Aspects of the theory of risk-bearing* 1965; (Yrjo Jahnsson Saatio, Helsinki).
- [17] Pratt, J. Risk aversion in the small and in the large. *Econometrica* 1964; **32**: 122–136.
- [18] Nord, E. Methods for quality adjustment of life years. *Social Science and Medicine* 1992; **34** (5): 59–69.
- [19] Schapira, M.M., Matchar, D.B. and Young, M.J. The effectiveness of Ovarian Cancer Screening. *Annals of Internal Medicine* 1993; **118** (11): 838–843.

Table 1: Social value of health programs

<u>Health gain, t (yrs)</u>	<u>Number of patients, p</u>	<u>Social value⁽¹⁾, p^*</u> <i>(t-Student⁽²⁾)</i>
10	10	10
20	5	9.23 (-1.90)
5	20	8.93 (-2.17)
2	50	7.49 (-2.99)
50	2	7.40 (-5.39)
1	100	6.74 (-3.49)

(1) *Number of patients who would have to receive a 10 life-year increase in order that this program be indifferent to the (t,p) program.*

(2) *$H_0: p^*=10$; $H_1: p^*\neq 10$*
 $n=61$

Table 2: Ranking of health programs⁽¹⁾

$S - PTO$ <i>(Health gain, (yrs),patients)</i>	$S - DO$ <i>(Health gain, (yrs),patients)</i>
10 , 10	10 , 10
20 , 5	5 , 20
5 , 20	20 , 5
50 , 2	50 , 2
2 , 50	2 , 50
1 , 100	1 , 100

(1) *From more to less preferred (average). n=61*
KCC=0.86 ; SCC=0.94 ; Individual SCC(average)=0.81

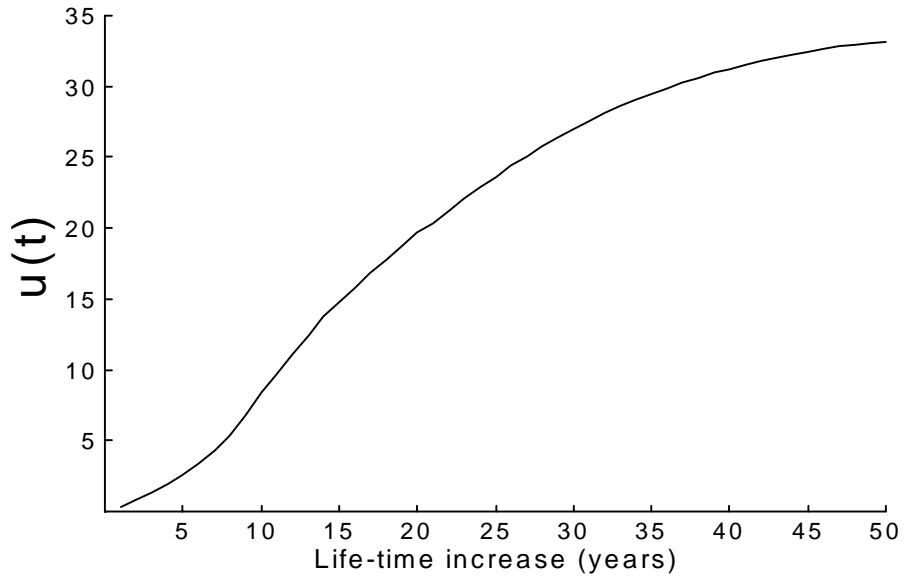


Figure 1: Social value of life-time increase

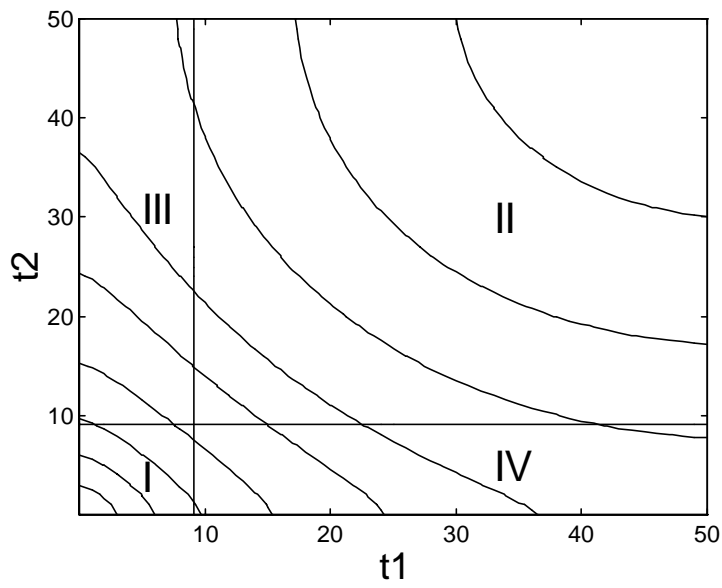


Figure 2: Indifference curves of the estimated SWF