Quality-adjusted life years (qalys) versus willingness to pay in matters of life and death

Yew-Kwang Ng
Monash University, Clayton, Victoria, Australia

It is an honour to be invited to contribute a paper in honour of Professor Clem Tisdell who is not only a prolific economist but has interests covering a wide range of topics of social significance. Similarly, this paper addresses an important issue that involves not just purely economic considerations – the issue of lives and deaths.

1. Introduction
In allocating scarce health-care resources between competing uses, difficult decisions have to be made. Recently, the concept of quality-adjusted life years (qalys) has been proposed as a measure of the benefits of prolonging lives [1]. In an interesting paper, Broome (1993) raised a “serious difficulty over using qalys to measure the goodness of alternative actions” (p. 161). A part from this, Broome, also made some important classifications and interesting discussion on qalys. In the present paper, it is argued that Broome’s “serious difficulty” can be overcome, at least in principle. A completely different difficulty, perhaps no less serious, on the use of qalys is also discussed. While we refer to qalys, the arguments have more general applicability to public resource allocation problems.

“{}There are two ways of bringing it about that good lives are lived – of producing qalys, that is. One is to prolong a person’s life or make it better. The other is to bring into existence a new life... Should one give equal value of qalys brought about by either method?” (Broome, 1993, pp. 161-2). The traditional procedure is not to count the qalys brought about by creating a new life at all. However, a practical and a theoretical problem exists for this solution, as discussed by Broome and some philosophers (e.g. Kuhse and Singer 1988; Parfit 1984). The practical problem refers to the difficulty of deciding about when a person comes into existence. This involves the distinction between a foetus, a prematurely-born baby, a normal baby, and an adult, etc. Broome thinks that it may be possible to find a philosophically defensible way around this practical problem but suspects that “the theoretical problem is insoluble” and that “the traditional procedure has no sound justification”. However, Broome also finds unacceptable the classical utilitarian alternative of counting all the qalys including those enjoyed by people brought into existence by an action.
In the next section, I support Broome, Parfit, and others in rejecting the acceptability of the traditional procedure. However, I also argue for the classical utilitarian position but with a modification to account for our (i.e. the existing people’s) partiality towards our own welfare. In Section 3, some elementary mathematics of some fundamental problems of ethics is presented that is relevant here as well as useful for clarifying other ethical issues. In Section 4, I raise a completely different difficulty for using qalys, one which involves Pareto inefficiency in comparison to the use of some method based on the willingness to pay.

2. The proper treatment of prospective individuals

The “theoretical” problem of the traditional procedure (which does not count the qalys of new people created) concerns “prospective” individuals who are not yet in existence but may be brought into existence under some alternative(s) under consideration. Considerations of prospective individuals lead to perplexing problems, as ably discussed in Parfit (1984) and Broome (1985, 1996). The main difficulty is this. If the welfare of prospective individuals is allowed into the objective function such as in the maximisation of total welfare (= average welfare times number of population), we can be led, under compelling conditions, to the so-called “repugnant conclusion” that a situation of no matter how small (but positive) average welfare (provided the size of the population is

A view over St. Peter’s which Clem Tisdell visited in 1965
large enough) is preferable to our current (or any other given) situation. On the other hand, if the welfare of prospective individuals is not included in the objective function, such as in the maximisation of average welfare of existing people, a compelling principle (the Mere Addition Principle to be discussed presently) is violated and the extinction paradox cannot be avoided (Ng, 1986), as shown below.

Suppose that releasing a certain chemical into the atmosphere would make everyone on earth healthier, happier, but sterile (but sexually still active, if not more so). Everyone regards the gain as more than the loss of the ability to have children. So everyone would be better off. But the release of the chemical would make mankind become extinct after existing people die off. Assuming that animal welfare is unaffected and that future people can expect to enjoy life of a significantly high positive welfare, the release of the chemical increases the welfare of existing people but is in my view totally unacceptable as it denies the welfare opportunity for prospective individuals.

The maximisation of the average welfare of existing people also violates the compelling Mere Addition Principle. This principle says that, if none of the existing individuals is made worse off, the mere addition of some happy individuals who, after having been brought into existence, would prefer to live, is a superior situation. This principle is something of a natural extension of the Pareto Principle to the case of a variable set of individuals and hence is extremely compelling. Let us use (100, 90) to indicate the situation of two individuals with 100 units of welfare for the first and 90 for the second. Then the Mere Addition Principle dictates that \( B = (100, 100, 60, 60) \) should be preferred to \( A = (100, 100) \). Non-antiegalitarianism \(^2\) (also compelling) dictates the preference of \( C = (90, 90, 90, 90) \) to \( B \). In fact \( C \) is preferable to \( B \) by either the total welfare principle or the average welfare principle. It is then not difficult to see that \((1, 1, \ldots, 1)\) for more than 200 individuals) is preferable to \((100, 100)\).

More generally, we have the repugnant conclusion. Thus, it can be shown that, given the compelling non-antiegalitarianism, the acceptance of the Mere Addition Principle entails the repugnant conclusion. Is the repugnant conclusion really repugnant?

For comparisons of ethical desirability, one should take an impartial view, abstracting from possible partiality towards one's self interest by adopting the viewpoint of an impartial outside ethical observer. In this perspective, the repugnant conclusion is not only not repugnant but is really compelling. \( C = (90, 90, 90, 90) \) is compellingly strictly preferable to \( B = (100, 100, 60, 60) \) which is compellingly strictly preferable to \( A = (100, 100) \). Thus the preference of \( C \) over \( A \) is also compelling. Similar logic leads us to the compelling preference of \( D = (1, 1, \ldots, 1) \) (for 500 individuals) over \( A \). The reason why most people find the preference of \( D \) over \( A \) repugnant may be due to their identification, consciously or subconsciously, with the situation \( A \). Looking from the viewpoint of the two individuals in \( A \), then obviously \( D \) is a repugnant result. The gain in one unit of welfare of each of the other 498 individuals in \( D \) does not
It might be asserted that one may prefer a situation (say $A'$ with 100 units of welfare for each of the five billion people) to another ($D'$ with one unit of welfare for 1,000 billion people) even if the latter involves higher total welfare provided the average welfare of the former is higher. There are at least two problems with such a position (held by almost all economists who have discussed the issue of optimum population; the only exceptions I know of are Dasgupta, 1969, Meade, 1955, and Ng, 1986). The first is that it violates, as already shown above, either the Mere Addition Principle or non-antiegalitarianism, both compelling. Another problem is that most people who prefer $A'$ to $D'$ and have declared adherence to average welfare maximisation also reject the logical consequence of preferring a utility monster (a situation $M$ of a single individual with welfare $= 1,015$) over $A'$. That the preference is seriously affected by the partiality to our current situation or something closer to our current situation ($A'$) seems clear. If $A'$ can be said to be better than $M$ (despite a lower average and total welfare in $A'$ due to more people enjoying positive welfare, that $D'$ should be preferable to $A'$ is much more compelling. If we can rid ourselves of all partiality, it may not be impossible to see that the impersonal total welfare maximisation principle (which meets the Mere Addition Principle and non-antiegalitarianism but entails the repugnant conclusion as well as the hypothetical, but not practical, possibility of a utility monster) is really compelling for comparing the pure ethical desirability of different situations involving different set and number of individuals.

However, the non-relevance of our partiality towards our own welfare to purely ethical comparisons does not mean its non-relevance for actual decisions. For simplicity in illustrating the relevant point, consider hypothetical situations. Suppose our current situation is $A'$. If by pushing a button, the world will suddenly become $D'$, I personally strongly believe that $D'$ is a better world than $A'$. But I will not push the button. This conflict between ideal ethics and practical decisions also applies to issues not involving a variable set of population. Consider a fixed set of four persons with the current situation being $E = (10, 10, 6, 6)$ where the first individual is me. If pushing a button will change the situation into $F = (7, 7, 10, 10)$, I will not push the button since I would be personally worse off despite that I have to agree at the impartial ethical level that $F$ clearly is a better situation than $E$ from an overall point of view. I see absolutely no inconsistency in admitting the social preferability of $F$ over $E$ and my private decision not to change from $E$ to $F$. Similarly, mankind (people existing now) can adhere to the impersonal total welfare principle at the level of ethical desirability for choices involving different sets of future population while refusing to adopt policies leading to the repugnant conclusion now. This distinction between ideal ethics and practical decisions influenced by partial interests thus effectively resolves paradoxes or difficulties besetting philosophers and economists for a long time concerning issues involving different sets of people.
The distinction between ideal ethics and practical decisions does not mean that principles (such as the impersonal total welfare principle) formulated at the level of ideal ethics have no practical relevance. Many of our choices may not mainly affect us (people now existing) but affect people in the distant future. For such decisions, we have no reason to be partial and should thus choose on the ground of ideal ethics. Models of interlocking generations with each generation being concerned with the welfare of its next generation may yield results that the welfare of people in the distant future still affects the welfare of the current generation through the interlocking effect. However, in practice, the interlocking is important only for objective effects (e.g. parents directly affected by children's actions and achievements) that are severed after more than one or two generations; the subjective interlocking that goes on forever is negligible at best. Thus, the welfare of the current generation is significantly affected by the objective circumstances perceived to affect the welfare of the next one or two generations but not appreciably by the circumstances in the distant future which affect the welfare of the 6th generation which affect the welfare of the 5th generation and so on.

3. Mathematical ethics
To highlight the distinction between self and others, one may abstract away issues of family obligations, friends, fellow countrymen, and so on. The fact that my private choice is mainly concerned with my own welfare does not mean that the welfare of others does not affect my preference or choice. The welfare of others may affect my welfare. But even if it doesn't, my preference may still be affected to some extent through the direct effect (over and above and separate from any effect through the effect on my own welfare) on my preference. To prove that such an effect exists, consider the following hypothetical choices. Readers (including most economists I have discussed the issue with) doubting the existence of this direct effect should put themselves into my situation described below.

Suppose I had to press one of two buttons A and B within a second. If I didn't, the world would be transformed into hell where everyone including myself would suffer enormous pain without being able to commit suicide. If I pressed button A, I would enjoy a blissful state of 100 billion units of welfare and everyone else on Earth would be condemned to a hell-like state (separated from me) of minus a million units of welfare. If I pressed button B, I would enjoy an almost blissful state of 90 billion units of welfare and everyone else (also separated from me) would enjoy positive welfare of one billion units each. Suppose also that within the second I had to push the button, I was too preoccupied with the decision such that my welfare within that second was zero whichever button I pressed and that after pressing the button I would lose memory and would not suffer regret or guilt feelings. Then, by construction, I would increase my welfare by pressing button A. However, I would not hesitate in choosing to press button B. I believe most people faced with the same situation would press button B. For those who would hesitate, change 90 billion
Quality-adjusted life years (qalys)

\[ U^s = F \{ W^x, W^y, ..., W^z \} \]

The welfare of others \( W^0 \) may not only affect my preference (represented by my utility function \( U^s \)) through its effect on my welfare \( W^s \) but also directly.

The first term on the right-hand side is the indirect effect and the second term is the direct effect.

\[ \frac{\partial U^s}{\partial W^0} = \frac{\partial F}{\partial W^0} \frac{\partial W^s}{\partial W^0} + \frac{\partial F}{\partial W^0} \]

If we interpret \( W^0 \) as the aggregate welfare instead of the average welfare of other individuals, then \( \frac{\partial F}{\partial W^0} \) is much larger than \( \frac{\partial F}{\partial W^0} \). Fundamental immorality may be defined by \( \frac{\partial F}{\partial W^0} = 0 \). (Called “fundamental” because it can be convincingly argued that moral principles should ultimately be based on individual welfare; see Ng, 1990.) A negative \( \frac{\partial F}{\partial W^0} \) signifies fundamental malice. Perfect morality requires \( \frac{\partial F}{\partial W^s} = \frac{\partial F}{\partial W^0} \). For completeness, \( \frac{\partial F}{\partial W^s} < \frac{\partial F}{\partial W^0} \) may be called excessive altruism which probably does not exist (realising that the altruism of mothers towards their children is largely effected through the indirect effect).

Coming back to the issue of different sets of population, we may interpret \( S \) as existing people or the current generation and 0 as prospective people or the future generations. Then, perfect morality dictates the impersonal total principle requiring \( \frac{\partial F}{\partial W^s} = \frac{\partial F}{\partial W^0} \). However, just as the choices of a typical person exhibits some but far from perfect morality, the current generation follows \( \frac{\partial F}{\partial W^s} > \frac{\partial F}{\partial W^0} > 0 \). In terms of the use of qalys, this means that the qalys of existing people are given full weights while those of prospective people are given only partial weights.

There is still the practical problem of determining who are “existing people”, the issue of unborn foetuses and premature babies. I do not wish to go into the full complication of this practical problem here. Suffice it to say that the principle of using different weights may be applied to this issue as well.

4. Qalys vs. willingness to pay

While I see no fundamental difficulty in using qalys due to the problem of prospective individuals, I have a completely different difficulty.

Suppose that we can only treat either one of the following at the same cost (e.g. only one machine is available for the purpose):

- A young person of 20 who is willing to pay a maximum of $50,000 for the treatment and (given the treatment) is willing to accept a minimum of
$80,000 to forgo the treatment which will extend his/her life by ten qalys; [3];

- A middle-age person of 50 who is willing to pay a maximum of $100,000 for the treatment and (given the treatment) is willing to accept a minimum of $150,000 to forgo the treatment which will extend his/her life by five qalys.

If we choose in terms of qalys, we will decide to treat the young person. However, if we treat the middle-age person instead and get her to pay $90,000 which is given to the young person, both persons will be made better off (in comparison to the situation where the young person is treated and no payment is involved) and no one will be made worse off. It is thus Pareto superior to treat the middle-aged instead of the young despite the indication to the contrary in terms of qalys.

It may be objected that if we allocate medical resources in accordance to the willingness to pay and accept compensation, the allocation will favour the rich and is hence unacceptable. Three points may be made in response to this objection.

First, the reason the middle-age person is willing to pay/accept amounts much higher than the young person despite a lower qalys involved need not be due to the fact that the middle-age person is richer than the young person. There are at least two other reasons that may account for the contrasting willingness to pay/accept. First, there is simply the individual differences in preferences with respect to current wealth versus future qalys, similar to individual differences with respect to specific goods, even with the same level of income. Then there is the difference due to the different stages of life cycle the two persons are in. As shown in Ng (1992), even holding both the utility function and the overall life-time income the same, and even assuming that a young person has no difficulty in borrowing against his/her future incomes, a 60 year old person may be willing to pay many times more in current dollars to avoid the same probability of death than a much younger person. The reason for this divergence between the utility (or qalys) and dollar values of life may be briefly explained.

At a positive real rate of interest (say 3-5 per cent), a dollar at the same point in time is worth much more to a young person than to an old person since it is received at an earlier stage in the life-cycle and can be used to accumulate for a longer period for the young person. The dollar value of life (or health, safety, etc.) equals the utility (or qalys) value of life divided by the marginal utility of a dollar. (This is similar to: the dollar value of an apple = the marginal utility of an apple divided by the marginal utility of a dollar.) If the marginal utility of a dollar is much higher for a young person, his/her dollar value or willingness to pay/accept may be lower than an older person despite a higher utility or qalys value.

Since the relative willingness to pay/accept may diverge from the relative values of qalys due to differences in age and preference (on top of income), using
qalys in making decisions involves forgoing Pareto improvements (as already illustrated above) even for cases where the rich-poor issue is not involved. It may be thought that qalys is certainly more fundamental than money and the former should be used instead of the latter where there is a divergence. However, the crux of the matter is that money is directly transferable at an almost one to one ratio (i.e. the costs of transfer are negligible and the divergence between willingness to pay and to accept is usually small for marginal changes) while qalys, utility, welfare or other fundamental and subjective measures of gains and losses are not transferable. For choices that harm one person and benefit another, we have to rely on transfer payments or compensation to make it a Pareto improvement. Thus, we should use the transferable indication of the willingness to pay/accept instead of other, even though more fundamental, measures. This point will be discussed further on another occasion. Now, we turn to our second and third points in response to the objection of using willingness to pay/accept.

Secondly, if the relevant compensation is only hypothetical, such as in the application of the Kaldor or Hicks compensation test, the objection on the ground of income distribution may be relevant. However, if the compensation is actually effected, then the relevant choice (as in the above example of treating the middle-aged instead of the young person) is actually a Pareto improvement. Even if the middle-age person is rich and the young person poor, it should be a social improvement to make both a rich and a poor person better off and no one else worse off.

Thirdly, consider the case where the compensation is not actually paid. It may appear obvious that we should choose in terms of qalys instead of willingness to pay/accept. Looking only at the direct effect of a particular instance of choice, it seems clear that the choice in accordance with qalys makes more sense than the choice in accordance with the willingness to pay/accept if compensation is not paid, since qalys is more fundamental than money, especially if the divergence between qalys and money is due to the rich-poor differentiation. However, if we take into account the indirect disincentive effects of equality-oriented policies, the picture is quite different. As shown in Ng (1984) the objective of achieving more equality is better pursued by a general tax-transfer system instead of through specific purely equality-oriented policies. The former has disincentive effects but the latter have specific distortive effects on top of the similar disincentive effects. This is so even if second-best complications are considered. If helping the sick causes less efficiency losses due to disincentives and moral hazard than the general tax-transfer system, it may be desirable to use sickness as an index of transfer such that public subsidies to health care may be desirable apart from the issue of external benefits. However, such a subsidy should only differentiate between the sick and the healthy and should not differentiate between the sick poor and the sick rich. (Otherwise it is Pareto-dominated by a policy dispensing with the latter type of differentiation, according to the proposition proved in Ng (1984).)
Thus, the objection to the use of willingness to pay/accept instead of qalys based on the rich-poor issue does not apply.

It is possible that the diamond effect (valuing goods and services for their values on top of their intrinsic consumption effects) may apply to some items of health care for the rich. If so, they should be charged higher prices. But this is based on a quite different principle of efficiency, not on the objective of equality. (See Ng (1987) on the diamond effect where taxes may impose not only no excess burden, but no burden at all.)

If the above point regarding disincentive effects is fully taken into account, one will avoid the mistake of regarding the issue of using qalys (hence preferring to treat the young despite lower willingness to pay) and that of transfer as two separate problems. As one commentator puts it, “Whether the money is transferred or not, it is better to treat the younger person than the older. The transfer of money is a red herring”. This quoted argument ignores the existence of disincentive effects. In fact, in the absence of disincentive effects and other costs of transfer, transfer should be made until willingness to pay and qalys agree. Then one could use either one. But this is Disneyland. In the real world, transferring the money from the old to the young (or from the rich to the poor) involves significant disincentive effects. However, if money is received as payment for treatment voluntarily sought, less or even no disincentive effects are caused. Thus the issue of transfer is not a red herring.

If we (i.e. the society) want to help the poor or the young, we should, in the absence of efficiency factors, such as external effects and ignorance, do so by the method of tax/transfer. For whatever extent of transfer we settle for, trading off costs and benefits, it is better to use the Pareto efficient method (willingness to pay) in resource allocation.

If goods are to fall, once only, from Heaven to householders, I wish that they are received by the poor. (Such acts of God will be perceived to be purely random and have no disincentive effects.) In this sense, a dollar (of goods) to the poor is valued higher than to the rich. But if treating a dollar to the poor as more important is used as a policy principle in specific resource allocation problems, it will produce disincentive effects in addition to its distortive effects and hence it is inferior to confining helping the poor through tax/transfer and using the principle of “a dollar is a dollar” in resource allocation.

The arguments for using willingness to pay and treating a dollar as a dollar are based on efficiency consideration of achieving whatever tradeoff between efficiency and equity. I am not arguing for more or for less transfer. Also, the arguments are not based on regarding the market as a morally free zone, requiring us to accept whatever outcomes. (See Gauthier (1986, Ch. 4) and a critic by Hausman, (1989).)

5. Concluding remarks
In this paper, we discuss the acceptability of using qalys in decisions affecting the length and quality of lives of people and of the existence of prospective individuals. In Section 2, the traditional procedure of ignoring prospective individuals...
individuals is dismissed as unacceptable. Instead, the classical utilitarian position modified to take account of our (i.e. existing people's) partiality is defended. In Section 4, the use of qalys is shown to involve Pareto inefficiency in comparison to the use of willingness to pay/accept. This latter usage is defended against the common objection based on the rich-poor issue.

However, I realise that it is practically and politically difficult to adopt the criterion of willingness to pay/accept in health care and other matters of life and death. First, due to the possible intertwining of the effects of external benefits and the diamond effects, it may be efficient to subsidise the health care of the poor much more than that of the rich. Secondly, due to the procedural dispreference (which is partly based on the ignorance of the efficiency issues involved; see Ng (1988)) for the use of willingness to pay/accept, its use in matter of life and death may not be politically acceptable. However, as economists, we have the role of clarifying the relevant issues for ourselves and educating the public. Perhaps, when we are defunct, some practical effects may then be felt, as Keynes believed.

### Notes
2. Non-antiegalitarianism is defined thus: If alternative B has the same set of individuals as in alternative A, with all individuals in B enjoying the same level of welfare as each other, and with a higher total welfare than A, then other things being equal, alternative B must be regarded as better than A.
3. On the proper treatment where the willingness to pay diverges greatly from willingness to accept (making $\Sigma CV$ and $\Sigma EV$ of different signs), see Ng (1979/1983, Appendix 4A).

### References
Ng, Y.-K. (1992), "The older the more valuable: divergence between utility and dollar values of life as one ages", Journal of Economics (Zeitschrift für Nationalökonomie), Vol. 55, pp. 1-16.