version 1(5), 2 Dec. '98
5100 words notes exclusive
HESG/Birmingham, Jan '99
work in progress
comments very welcome
do not quote without permission

Does CBA really have a better welfare economic foundation? Going back to basics, and re-examining welfarism

Aki Tsuchiya, at7@york.ac.uk
Centre for Health Economics, University of York

1. Introduction

Economic evaluation of health care with the purpose of providing information for resource allocation offers, in one way or other, analysis of inputs to and outputs form given health care interventions, and compares these across different interventions. While inputs are almost always represented in monetary terms, outputs can be various, ranging for example from a change in some clinical indicator, difference in remission rates, numbers of lives saved, to crude or quality-adjusted life-years gained, all of which are non-monetary representations, and finally all of these and anything else that may be of relevance can, in theory, be expressed in monetary terms. When outputs are expressed in monetary terms, an economic evaluation is called a CBA (cost-benefit analysis), and when they are not, it is referred to as a CEA (cost-effectiveness analysis). When a CEA employs QALYs (quality-adjusted life-years) as the output measure, it is sometimes referred to as a CUA (cost utility analysis). Nevertheless, hereafter, this paper uses the term CEA as a shorthand for CEA with QALYs.

Those in or familiar with the medical community, faced with its reluctance to assigning monetary values to life and morbidity, have long advocated the use of CEA over CBA. On the other hand, there has been criticisms from the economics field arguing that WTP(willingness-to-pay) and CBA have better welfare economic foundations compared to QALYs and CEA, and therefore CBA is the theoretically correct way ahead.

The purpose of this paper is to offer different interpretations of CBA and CEA, and in so doing, re-examine new welfare economics, the potential compensation criterion, and welfarism. "Health

gain" here is not limited to the alleviation of non-fatal conditions, but will also include life-saving. Section 2 lays out a description of a "value free" way of evaluating health change in monetary terms at the individual level, and another of representing improvement at the collective level. The question is, how to use information on the former to determine whether the latter is taking place. Section 3 first demonstrates that this is subject to serious difficulty, and then discuss different approaches that can be interpreted as attempts to circumvent this difficulty[1]. The 4th and 5th sections deal with approaches that divert from the initially intended value-freeness.

2. A "theoretically correct" concept of the monetary value of health

Current mainstream (or, textbook) economic theory holds that (a) utility is a subjective phenomenon and as such is not measurable or comparable in any objective and scientific manner, and (b) social welfare should be based on individual utility and nothing else. The most important implication of (a) is that building welfare economic theory on fully measurable and comparable utility is unacceptable. Point (b) is often referred to as "welfarism"[2]. The school, so-called *new* welfare economics, advocates the use of ordinal utilities with no interpersonal comparisons permitted, and offer theoretical tools with which to analyse social welfare within welfarism.

Monetary value of a given health change with reference to a given initial state can be conceptualised as a CV (compensating variation) or an EV (equivalent variation), and in practice be elicited as a WTP (willingness to pay) or a WTA (willingness to accept)[3]. For example, Figure 1, adopted from Johansson(1995), illustrates the WTP for a given improvement in health (i.e. a CV). The axes represent the level of health of an individual and his/her income. Starting from point *e*, an improvement in the level of health (from H_0 to H_1), with no change in either income or relevant prices, is represented by a move to point *f*. CV corresponding to this change is defined as the vertical difference between points *f* and *g*: i.e. the amount of money, if taken away from the individual after the change, will bring him/her back to the original utility level (U_0). A typical interview question designed to elicit this amount will ask:

---

[1] The objective is *not* to claim that these different approaches have been explored with the *intention* of circumventing the particular problem studied here. Further, the paper does not attempt to do justice to the *historical* development of the field.

[2] This is an unhappy terminology, since it is named after the defined element or dependent variable (*viz.* social welfare) rather than the defining element or independent variable (*viz.* utility) of the relevant relationship. Something like "utilitism" may have been a more straightforward term.

[3] It is complicated, but while CV and EV are distinguished from each other with reference to the relative prices either posterior or prior to the change in question, WTP and WTA are distinguished from each other with respect to the direction of the money flow. See Box 7.8, Drummond et al.,1997. Note also that, regarding non-convergence, the CV-EV literature is concerned with income effect while the WTP-WTA literature is more concerned with psychological factors. This seems to reflect the fact that while CV and EV are theoretical constructs to conceptualise the magnitude of a given change, WTP and WTA are practical tools with which to express non-marketed values in monetary terms.

"What is the maximum amount of money you are willing to pay in order to enjoy an increase in the level of your health from H_0 to H_1?" (Q1)

Let us assume that the replies thus obtained are reliable and valid. Note that this money-equivalent is likely to depend on the level of y, because the only restrictions on the nature of individual utilities are that they are increasing in the levels of income and of health, and that marginal rates of substitution decrease[4].

The next stage is to aggregate these monetary evaluations of different individuals thus obtained and to determine whether for example a proposed change in health care resource allocation will be preferable at the social level. The standard procedure is to appeal to the potential compensation criterion (PCC hereafter). PCC claims that, if the winners from a programme are *possible* to compensate the losers for their losses so that nobody was worse off and at least one person was better off with the compensation, then a programme is worth adopting *without* the compensation actually taking place[5]. Note that: (a) the criterion is applicable to situations that *fail the Pareto criterion*[6]; (b) the criterion operates *without* stepping into the realm of measurable and/or comparable utilities; and (c) the proposed output-mix is evaluated without regard to how the outputs are distributed amongst the individuals. This last point means that PCC concerns efficiency only and distribution is to be accounted for elsewhere[7].

This is illustrated in Figure 2. The Figure depicts a simplified economy with two goods (x_1 and x_2), the present output-mix (E), the proposed alternative output-mix (F), and a community indifference curve[8] (SS), drawn with respect to the present distribution (between two individuals i

[4] This is consistent with what is actually observed in the WTP literature: *viz.*, WTP for an increase in life expectancy or a decrease in risk of death, is known to have a positive correlation with a respondent's income.

[5] There are two versions: the Kaldor criterion, which concerns whether the winners could compensate the losers under the prices *after* the change, and the Hicks criterion, which concerns whether the losers(-to-be) could bribe the winners(-to-be) under the prices *before* the change. The difference corresponds to the difference between CV and EV.

[6] To refer to this potential compensation criteria simply as "Paretian" is, in this sense, highly misleading. Obviously, whether the relevant criterion of choice is the actual existence of a Pareto improvement or the existence of a mere potential of it will be crucial to a pure Paretian, because the whole essence of the Pareto criterion lies in not accepting anybody to be a looser.

[7] In order to justify this unconcern for distribution, it is assumed that efficiency and equity can be dealt with separately. This obviously is a highly controversial claim, but is not discussed here. The purpose of this paper is not to examine the legitimacy of PCC, which is a matter of judgement, but to explore to what extent CBA can be faithful to its purported foundations. Also, see the comments in Blaug,1992 on the position of new welfare economics: judging whether actual compensation *should* be paid in order to bring forth a real Pareto improvement is value laden, "but no value judgement whatsoever was involved if the economist merely described a change as a potential Pareto improvement" (p.123), and thus a "value-free" new welfare economics.

[8] The community indifference curve is also called the Scitovsky frontier. The curve, in the two person case, is drawn by manipulating the Edgeworth box, which has points O and E as the two origins. A community indifference curve is defined for a given particular distribution of goods along the contract

and j) of the output-mix represented by point E. The individuals do not appear explicitly in the diagram, but the curvatures of their personal indifference curves have contributed to the curvature of SS. The SS curve ensures that any given point north-east to it can, with appropriate distribution, yield at least as much utility as any point on the curve for the individuals concerned: i.e. if point F lies inside the SS contour, there is potential for compensation. Therefore, if PCC is to be the underlying welfare economic foundation for CBA, then the latter should be identifying whether the point representing the proposed change lies inside the relevant community indifference curve or not. The next section will study how this is not the case.

3. Adopting the concept in real world analyses

3.1. The sign of ΣCV

In a CBA, not only WTP for a particular health improvement, but also monetary values (i.e. competitive prices) of inputs for its production are represented as (negative) CVs of applicable individuals, and all these are added together. It is an attempt to determine whether the monetary value of all outputs exceeds that of all inputs, and therefore, it is a test concerning whether the sum of all relevant CVs is positive or not. The rule of thumb at the base of CBA is that if $\Sigma_i CV_i > 0$, *i.e.*, if the simple sum of CVs of all people concerned is positive, then there is to be a net gain to the overall community. There being a "net gain" is to coincide with the satisfaction of PCC. Nevertheless, as is shown below, a straightforward interpretation of PCC is not equivalent to ΣCV>0 (Boadway,1974; also see Boadway,Bruce,1984, Okuno,Suzumura,1988).

Representing ΣCV in the diagram involves drawing straight lines, with a slope corresponding to the relevant marginal rate of substitution between the two goods. This is shown in Figure 3, where line MM has a slope representing the relative prices of the two goods under point F[9]: i.e. the line represents all combinations of goods x_1 and x_2 that are equivalent in monetary terms to point F, provided that none of the relevant prices are distorted. Given the relative prices, distance OM corresponds to the minimum aggregate income, in terms of the corresponding good, to maintain the level of utility at point F. Now, add line NN parallel to line MM and tangent to the community

---

curve. The personal indifference curve of the person whose indifference curve is drawn with respect to origin O is fixed, and the indifference curve space of the other person is shifted along the curve of the first, maintaining tangency, and keeping the corresponding axes parallel to each other. The resulting locus of origin E depicts the community indifference curve. One contour can be drawn for every possible distribution along the contract curve, intersecting each other at point E. For further details, see for example Boadway,1974, Boadway,Bruce,1984.

[9] The format here deals with the relative prices under point F, and thus concerns CV and the Kaldor version of PCC. For EV and the Hicks version of PCC, relative price at point E becomes the reference. Nevertheless, when relative prices are assumed to remain constant, CV and EV are equal to each other.

indifference curve (*SS*) for point *E*. The point of tangency is labelled *G*. Since distance *OM* (*ON*) along either axis represents the sum of income necessary to cover utility levels at point *F* (*G*), if *OM*>*ON*, then this means that the monetary amount that can, with appropriate distribution, ensure the individuals their utility levels at point *F* is larger than the corresponding amount for point *G*, which in tern can provide, with appropriate distribution, the same utility levels for each individual as point *E*. Hence, *OM*>*ON* is equivalent to $\Sigma CV>0$, and thus, if point *F* lies north-east to line *NN*, $\Sigma CV>0$ is satisfied. Clearly, the area north-east to curve *SS* and the area north-east to line *MM* are not equal, and therefore $\Sigma CV>0$ and PCC are not equivalent[10]. Under the assumptions employed so far, the former is a necessary, but an insufficient, condition for the latter: if $\Sigma CV \leq 0$, then a programme ought not be taken up, but $\Sigma CV>0$ alone cannot justify its implementation.

The two shortcomings of PCC for practical purposes[11] are that (a) observing the community indifference curves in the real world is not easy, if not practically impossible, and despite this, (b) assuming diminishing marginal rates of substitution for both individuals, we can say, without actually observing these curves, that they will be convex to the origin and therefore the $\Sigma CV>0$ rule, which implies straight lines, cannot test for PCC. The next section will show how each of these may be circumvented, but not without difficulties.

3.2. Relaxing the criteria

The concept of PCC explained above is sometimes referred to as "strong" PCC, because there is another ("weak") formulation of PCC. (The two terms were first introduced by Foster,1979, but the concepts themselves appear in Boadway,1974. Also see Boadway,Bruce,1884, Okuno, Suzumura,1988, Oka,1997.)

The purpose of PCC is to see whether or not there is potential under point *F* to guarantee each individual his/her utility level achieved under point *E*. In examining for this, "strong" PCC considers different distributions, between the two people, of the *specific output-mix at point F*; whereas "weak" PCC allows *re*distribution between the individuals in terms of purchasing power and thus to *move along line MM* to explore the potential for compensation. This obviously

---

[10] This relationship is sometimes referred to as the "Boadway paradox", which is another misnomer. As long as the potential for compensation is the *criterion* and the reference to ΣCV a (possible) *test* for it, all this amounts to is that the test is a failure in that it does not test what it is supposed to test.

[11] A more abstract and well known shortcoming of PCC is that under certain conditions, PCC may lead to paradoxical conclusions (*viz*., a move from point E to point F is preferable under point E, and then the move back form point F to point E becomes preferable under point F). This happens because, unlike individual indifference curves, community indifference curves for a given output-mix may intersect with each other. Without a suitable means to operationalise PCC, this remains a purely theoretical problem. When, in order to make use of PCC in the real world, either PCC is relaxed (as in 3.2) or assumptions

requires that goods can (potentially) be imported from the outside world beyond our two-person economy[12]. Now, while "strong" PCC requires that *point F* lies to the north-east of the *SS* curve for point *E*, "weak" PCC requires that *line MM* lies to (or has a section that passes through) the north-east of the *SS* curve drawn for point *E*. Here, the inconvenience of the community indifference curve being unobservable is, in effect, overcome, and $\Sigma CV>0$ is equivalent to weak PCC thus formulated. Also note that neither cardinality or comparability of individual utilities is required. To see that weak PCC does not necessarily imply the satisfaction of strong PCC, Figure 4 gives an example. Since line *MM* has a segment that lies to the north-east of curve SS, weak PCC is satisfied, but since point F now lies on the west of curve SS, strong PCC is not satisfied.

While strong PCC deals with a move from one output-mix (point *E*) to another output-mix (*F*) and with the potential for compensation there, weak PCC involves yet another potential of moving to an unspecified output-mix (say, *H* in Figure 4) in order to satisfy the originally required potential for compensation. The justification for this extension is that: *F* is interchangeable with *H*, and *H* is more preferable than *E*, therefore *F* is more preferable than *E*.

Nevertheless, there are two serious caveats to note. The first is that being interchangeable in monetary terms does not guarantee indifference in preference terms, and therefore the syllogism is at fault. In Figure 4, point *F* lies outside contour *SS* and in real terms *is inferior* to point *E*. The second is that if we are to determine the desirability of a move from one specific situation to another specific situation, recourse to some unidentified intermediary output-mix seems somewhat illusive. These two points put together may well imply that we are to be satisfied with any given inefficient situation because, potentially, it is as good as being on any particular point on the production possibility frontier[13].

3.3. Tightening the assumptions

The above subsection has shown that circumventing the first of the two shortcomings, listed at the end of subsection 3.2 above, by weakening PCC is not without difficulties. This subsection discusses the feasibility of dealing with the second shortcoming listed.

---

regarding individual utilities are tightened (3.3), the paradox disappears.

[12] A more conventional formulation of weak PCC is to allow for moves along the relevant production possibility frontier through point *F*. This is not employed here because (a) it leads to a less attractive conclusion where $\Sigma CV>0$ is neither sufficient nor necessary for the satisfaction of weak PCC, and (b) the move should not have to be restricted to those along the PPF, since it is to be merely potential.

[13] There is an argument in Boadway,Bruce(1984: pp.96-7) that gives superiority to weak PCC over strong PCC. The assumption here seems to be that *both* the present *and* the proposed alternative situation are defined as output-*possibility*-mixes. This involves that difficulty that the community indifference curve cannot be drawn unless a specific output-mix and its distribution are given; or, in other words, weak PCC becomes independent from a specific initial state, while, the meaning of social desirability of a given change independent of initial states is unclear.

It is possible to ensure that ΣCV>0 is a test for strong PCC by letting curve *SS* coincide with line *MM*. This can be done by employing stringent assumptions so that at least one of the two individuals have straight indifference curves, which slope is equal to the relevant non-distorted prices. For at least one individual in *any* two individuals to satisfy this condition, all but one individual in a given community must, and therefore a more likely assumption may be to assume that *all* individuals have such preferences[14]. Assuming constant and equal marginal rates of substitution between health and income for all individuals means that everybody will be willing to exchange between these two at the same rates: i.e. a very rich person in very poor health will have the same marginal rate of substitution as a very healthy person with very little money; an assumption that is unlikely to be satisfied in the real world[15]. However, unless one agrees to accept weak PCC as the relevant criteria, this is the necessary assumption for the ΣCV>0 decision rule, and CBA, to make theoretical sense. Whichever alternative one holds, it seems right to say that, while the approach of new welfare economics may have started from "objective" and "scientific" premises by confining itself to ordinal and non-comparable utility, the practice of CBA involves stretching it to a point where the implications are no longer value-free.

There are two things to note. First, stringent these assumptions may be, there is still no explicit reference to measurability or comparability of individual utilities. Second, satisfaction of PCC does not concern the quantification (in cardinal measures) of different *degrees* by which given output-mixes may be superior or inferior to each other.

4. An alternative: social welfare functions (SWFs)

4.1. SWFs with monetary valuation of health

While the last approach above does not refer explicitly to cardinal and measurable utilities, its assumptions can be interpreted to imply them[16]. Firstly, regarding measurability, constant marginal rates of substitution implies constant marginal utilities regarding the goods in question (in our present case, health and income)[17], which requires that reference can be made to the actual size

[14] This may also be required for anonymity. Also, note that the distinction between strong and weak PCC becomes irrelevant under this approach.

[15] It also deviates from microeconomics in general, since the assumption of constant marginal utilities implies that optimisation problems will be insoluble.

[16] For example, Layard,Walters(1994) explicitly employs cardinal and comparable utility, with weak PCC, and an additive SWF. The argument for doing so is based on Arrow's Impossibility Theorem: i.e. in order to make normative statements at the collective level in a "democratic" manner, ordinal and incomparable utility is insufficient.

[17] The alternative will be to assume that, moving along a given personal indifference curve, marginal utility of either health or income decreases, while marginal utility of the other decreases too at the same rate so that marginal rates of substitution (which is a ratio between the two marginal utilities) remained

of each second derivative, which in tern implies cardinality. Secondly, regarding comparability, if, as assumed, all individuals concerned have identical cardinal utility functions, then the level of utility can be inferred from the levels of consumption, and thus compared across different individuals. Also note that constant marginal utilities further ensure that an individual's CV of, for example, £1,000 is derived from a utility change of a fixed magnitude irrespective of his/her level of income, and that identical utility functions mean that individual *i*'s CV of £1,000 is derived from the same magnitude of utility change as individual *j*'s CV of the same amount. These imply that CV_i is a cardinal and comparable measure of utility change.

While PCC per se is an ordinal criterion, this alternative approach allows to quantify how much better or worse a proposed output-mix is compared to another given output-mix, by reference to an appropriate SWF (social welfare function). The simplest SWF compatible with PCC is the so-called utilitarian SWF:

$$\Delta W = \Sigma_i \Delta v_i,$$
$$v_i = v(y_i, p),$$

where W stands for social welfare and v_i is the level of measurable and comparable indirect utility of individual i, which is a function of his/her income (y_i) and a given price vector (p), and which is the same across all individuals. Though the employment of cardinal and comparable utility is strictly against the tenets of new welfare economics, it is still welfarist because the only arguments in the first equation are utility of different individuals. Here, since marginal utility of income is assumed to be constant and equal for all, $\Sigma CV>0$ is equivalent to $\Sigma_i \Delta v_i>0$, which in tern implies an increase in social welfare, and the magnitude of ΣCV will represent the degree of this improvement in monetary terms. Under a given fixed price vector, change in social welfare becomes a function of aggregate individual income alone: in other words, these are the grounds for identifying growth in national income and/or GNP with change in social welfare.

An interesting thing to note is that, since all individuals are assumed to have identical utility functions, this approach does not allow for so-called preference subgrouping. To show the importance of this assumption, Figure 5 illustrates the case where two individuals each have different utility functions (with constant marginal utility). The vertical axis represents cardinal and comparable utility, and the horizontal axis represents income. Suppose that a proposed change will result in individual *i*'s income to fall from y_1 to y_0, and individual *j*'s income to rise from y_0 to y_2, while all other things are kept constant. Since the extent of CV_i (which is negative) is smaller than that of CV_j (which is positive), $\Sigma CV>0$. Now if, as is shown, individual utility

---

constant: an assumption that is even less attractive.

functions are not the same and u_i has a larger marginal utility than u_j, then this change will result in a decrease in W, defined as Σu. This demonstrates that the only way in which the sign of ΣCV can be referred to in accordance with the SWF defined above is by assuming that individual utilities have the stringent assumptions being discussed here, and therefore, for a given health change, a single value of CV will be used across the whole population[18].

Since in practice, WTP regarding a given improvement in health is known to have a positive correlation with a respondent's income and/or wealth, the subjective monetary value of health of the relatively rich (poor) is likely to be higher (lower) than this uniform value applied across the population. This may be seen by some as only fair. Nevertheless, though the general direction of giving smaller weights to the rich and larger weights to the poor may be acceptable, the actual extent of weighting is not easily observable and therefore not ideal.

It is sometimes suggested further that a more general form of the utilitarian SWF can be employed so that:

$$\Delta W = \Sigma_i a_i \Delta v_i,$$

where a_i represents the relative weight at which the utility of a given individual i contributes to social welfare: i.e. the SWF mentioned earlier corresponds to the case where a is the same for all individuals. In order to reflect this alternative SWF, not simple ΣCV, but weighted ΣCV needs to be used in CBA. The suggestion is an attempt to incorporate equity considerations explicitly within the CBA framework. It is based on the understanding that CV of the relatively poor is unduly affected by their restricted ability to pay, which is to be corrected by the introduction of distributional weights that correlate negatively with each individual's income/wealth. The theoretical difficulty of this approach is that the assumption of constant and equal marginal utility for all individuals already implies quasi-equity weights, and the additional treatment needs to be justified[19].

There are two things to note. Firstly, reference to the SWF above and reference to variance in WTP due to effect of income/wealth are incompatible. For instance, a study that appeals to the utilitarian SWF cannot at the same time report threshold levels of income/wealth over (under) which treating (not treating) an individual becomes cost -effective (-ineffective), because all individuals are to have identical CVs for a given change.

---

[18] If different monetary values for a given health improvement are employed across different individuals, what ΣCV>0 represents is the financial feasibility (in potential terms) of the change, and not the potential for compensation nor increase in the sum of individual utilities. There will be enough money round to fund the change, but increase in social welfare is not guaranteed.

[19] For instance, see Sugden,Williams(1978) for an argument that there is no scope for distributional weights within the realm of PCC (referred to, regrettably, as the "Paretian approach").

Secondly, the stringent and unlikely nature of the assumptions can be overcome if an alternative proxy for individual utility, where the same assumptions may be more appropriate, is employed. The QALY measure is an obvious candidate for this.

4.2. SWFs with non-monetary valuation of health

Let us introduce QALYs into the picture, by redefining the individual indirect utility function to include a separate argument representing the level of health, and maintaining the (unweighted) utilitarian SWF so that:

$$\Delta W = \Sigma_i \Delta v_i,$$
$$v_i = v\,(y_i, p, H_i),$$
$$H_i = q_i \,.\, L_i \qquad (-1 < q < 1),$$

where H_i indicates the level of health for individual i expressed in terms of QALYs, which is the product of the number of life years (L) and health-related quality of life (q)[20]. The latter is also referred to as the quality adjustment weight, and is derived by asking questions that, in essence, ask individuals:

> "How (un-)desirable would it be for you to be in health state X compared to full health for a given length of time?" (Q2)

Again, let us assume that the obtained replies are reliable and valid. The undesirability of a given health state relative to full health will be the preference-based quality adjustment weight q of this state. A welfarist interpretation of CEA requires that QALYs represent individual utility. The construction of the QALY is such that while q is preference-based, L is not, and therefore the requirement implies the assumption that all individuals have constant marginal utility from additional life years (at full health)[21]. Further, for a QALY of individual i to be comparable with a QALY of individual j, this constant marginal utility is to be identical across all.

Now, let us suppose that income (y_i) and prices (p) remain constant, so that v_i becomes a function of H_i alone[22]. Then, at the collective level, given a fixed budget for health care,

---

[20] This simple model can be extended to allow for time preference, and "health paths" with change over time in q, but incorporating these will not change the nature of the argument below.

[21] It may be worth pointing out that regarding QALYs, some authors restrict the use of the term "utility" to the case where q is derived from the standard gamble (SG) technique, but even when the value of q is based on expected utility theory, in order to refer to *QALYs* as utilities, one needs to assume risk neutrality over life years. Note that if risk neutrality is to be assumed, then the very argument that is seen by some to place SG superior to alternative techniques in the first place becomes mute, unless one assumes that people are risk neutral regarding the length of life, but risk averse regarding quality of life. This in tern will require different treatment of the same number of QALYs depending on different combinations of the length of life component and the quality of life component.

[22] In reality, as the value of H_i gets lower, it becomes increasingly more likely that y_i will be affected

allocating this so as to maximise the number of additional QALYs across the population will achieve the largest improvement in social welfare, other things being equal. This will be enabled by equalising marginal output from all available health care interventions, and therefore marginal cost/QALY analyses will play an important role. As this explanation indicates, the analysis now is clearly of the partial equilibrium kind, assuming that the allocation of resources outside the health care sector is given and constant[23].

A curious condition for this *at the individual level* is that, QALYs are assumed to be incommensurable with other goods and commodities that are traded for money. If not, there is no reason to allow an independent argument in the indirect utility function to represent them. Incommensurability implies that indifference curves, between QALYs and all other goods (represented by income) as in Figure 1, cannot be drawn, and therefore, now, PCC is implicitly abandoned. In other words, winners (i.e. those whose QALYs will increase as a result of a proposed resource allocation) cannot compensate losers (i.e. those whose QALYs will decrease) with money transfers *even in potential terms*. Nevertheless, since allocation of scarce resources will involve a cut-off point in terms of costs/QALYs, beyond which health care with positive-but-inefficient QALY gains will not be funded, the approach implicitly but inevitably results in commensurability at the collective level.

Note that this presentation of CEA does not satisfy the tenets of new welfare economics, nor PCC, but does fall within welfarism.

5. A re-interpretation of the non-monetary value of health

The approaches discussed so far are based on welfarism, which requires that social welfare be based on individual utility: i.e. health gains to individuals generate individual utility, which in tern add up to generate social welfare. The alternative interpretation now is to go beyond welfarism and to re-define social welfare as *an aggregation of individual well-being, which in tern is a function of* (amongst other things) *health gain to each individual*, where individual well-being is defined more broadly than individual utility. The main arguments for this diversion from welfarism stand on two claims:

---

unfavourably. In the ceteris paribus analysis, it is useful to assume that $v(\cdot)$ takes a positive value when $H_i > 0$ and is equal to zero when $H_i = 0$ (Johansson,1995).

[23] Going back to monetary representations of health change, CV assumes that there is no price distortion. Further, for CBA to dictate the so-called cut off point in an endogenous manner, general equilibrium needs to be satisfied so that the marginal value of money is the same across the economy; something that does not apply to the real world (see Olsen,Smith,1998). There should also be scope to discuss CBA in the context of partial equilibria (see Oka,1997).

(a) individuals are not always the best judge of their own welfare, and
(b) people's preference regarding their own welfare alone are not the only relevant things to account for in (health care) resource allocation.

The former point leads to the conclusion that health may be something individuals ought to be given regardless of how little an individual values it, and the second to the recognition that caring externalities and fairness may have an important role to play. Examples of such an approach in the health economics literature are "extra-welfarism" (Culyer,1989)[24], and the "decision-making approach" (Sugden,Williams,1978) [25]. Here, let us use "non-welfarism" to encompass these two and other similar deviations from pure welfarism (e.g. Williams,1972, Williams,1988, Nord,1994, Dolan,1998). Here is a simple illustration of the two approaches:

conventional model (welfarism):
health gains to individuals → individual utility → aggregation → social welfare

alternative model (non-welfarism):
health gains to individuals → individual well-being → aggregation → social welfare

Under non-welfarism, since the magnitude of individual well-being derived from health gains to individuals can be independent of how QALYs are evaluated by the corresponding individuals themselves, the assumption of constant and equal marginal utilities is replaced with that of constant and equal marginal impact to social welfare from each QALY. A simple non-welfarist SWF will look like this:

$$\Delta W = \Sigma_i \Delta s_i,$$
$$s_i = s\,(H_i).$$

Here, s_i represents the level of social value generated by the well-being of individual i. Since a typical non-welfarist approach in health economics is to set health as the maximand, and due to the ceteris paribus nature of CEA, all other things that may in theory enter into function s are here suppressed. That money and QALYs should be incommensurable at the individual level but commensurable at the collective level of CEAs can now be interpreted as reflecting a non-welfarist decision that the social value of a QALY should be independent from variation in individual patient wealth/income. The relevant indifference curves may be drawn, but will be seen as irrelevant to the context.

---

[24] "The other approach, which might be termed 'extra-welfarist', relaxes what its adherents see as an undue information restriction in welfarism so that other aspects of each social state are also embodied in the judgement." (Culyer,1989, p.36)

[25] "With this approach, then, cost-benefit analysis is seen as a process of appraising decision making problems in the light of objectives chosen by the decision-maker. We shall call this approach the *decision-making approach*". (Sugden,Williams,1978, p.91, emphasis original)

Furthermore, under this re-interpretation, weights can be incorporated which will readjust for the equality of marginal social impact of QALYs, to reflect unequal impact depending on to whom these QALYs are gained. Note that this equality/inequality is not imposed on how individuals evaluate marginal QALYs themselves, but reflects a judgement on how each QALY ought to be weighed in the social welfare calculus. For example, social welfare may be formulated as:

$W = (\Sigma_i b_i s_i^r)^{1/r}$, so that
$\Delta W = \Sigma_i c_i \Delta s_i$, where
$c_i = \partial W / \partial s_i$, and
$s_i = s(H_i)$.

Now, b_i represents the relative weight each individual is given in the calculus of social welfare: if the health of individual i is to be given equal social weight as any other, then $b = 1$ for all. r represents the curvature of the iso-social-welfare contour, which corresponds to how social welfare is affected by distribution of a given simple sum of social values: if social welfare is to be distribution-neutral, then $r = 1$. As the presence and absence of subscripts show, while the value of b_i represents each individual's attribute per se and can vary between individuals, the value of r represents an aversion to inequality of the society and there will be one value for all. c_i corresponds to the weights which reflect aversion to inequality in levels of health, and is a function of s_i, b_i, and r. Actual CEAs will take the form of $\Delta W = \Sigma_i c_i \Delta s_i$, and therefore c_i will be the actual weights employed.

On a more empirical level, the choice between welfarism and non-welfarism should affect the way in which health is valued. Nevertheless, little has been achieved to date concerning the actual values of any of the relevant weights or parameters.

While many CEAs seem to associate themselves with non-welfarism, most continue to rely on health state valuations derived in the manner represented as Q2 above. It is important to note that, employing the value of q thus obtained amounts to assuming $s_i = H_i$, which can only serve as a proxy, because it equates an individual's own preference over a given health gain with its social value. A non-welfarist approach requires that health changes are evaluated, without supposing that the respondents themselves are subject to the change in question: i.e. to elicit the social value s directly from a less personal perspective, for example, by asking:

> "How (un-)desirable would it be, for society as a whole, for somebody (not you) to be in health state X compared to full health for a given length of time?" (Q3)

Person Trade-Off (PTO) questions are used to elicit health state values from this perspective (see for example, Nord et al.,1993, Murray,1996, Richardson,Nord,1997). Nevertheless, PTO is expected to be affected by a respondent's concern for distribution, and therefore, the effect of r may

be confounding the obtained values which, theoretically, should correspond to s[26]. A better alternative may be to use time trade-off questions under the social perspective. Further, faced with Q3, people may want some attributes of this "somebody" in question to be specified. Exploring this line, by fixing the health state X and varying the description of the patient, the b weights may be obtained[27]. The r parameter can be elicited by posing questions on trade-off questions between overall output and its distribution[28].

Conclusion

The discussion has shown that, while there is a "theoretically correct" way within new welfare economics to derive monetary valuations of health, this was not a very practical tool when it came to the aggregation level, and it was found that the practice of CBA does not follow faithfully the vow of value-freeness. Five alternative interpretations of cost-benefit thinking were discussed, some of which allowed to reconcile practice and theory, either by relaxing the requirements of the theoretical criterion, or by posing more and more stringent assumptions on the nature of the value of health, while others diverted from the fundamental theory altogether. First, strong PCC was dropped, and then cardinality and comparability were introduced. While the introduction of QALYs alone under welfarism was not an attractive option, QALYs under non-welfarism seemed to offer an alternative comparable to CBA. While there will be those who continue to advocate the use of CBA over CEA, this should not be done by claiming theoretical superiority of the former. CBA is not a straightforward application of some well-established consistent theory, but a carefully contrived construction with a host of assumptions and value judgements attached. To some extent, the same applies to CEA under non-welfarism, but despite the considerable allusion to non-welfarism amongst CEAs, this approach at present seems relatively underdeveloped as an integrated system. Two important elements for empirical research for non-welfarism can be pointed out. The first is to study the effect of the social perspective on the valuation of health: i.e. to compare $v_i(H_i)$ and $s(H_i)$. The second is to investigate the scope for social weighting: i.e. the magnitude of coefficient b_i and parameter r.

---

[26] See Dolan,1998 for a critique of PTO.
[27] See Williams,1988, Wagstaff,1994 for a discussion. The latter is a reprint of an article published in 1991, with (what I understand to be) the correct form of the SWF misprinted in the original.
[28] See Williams,1997 for an example.

Acknowledgements

This paper was inspired while the author was working in a study group for WTP, financed by the Institute for Health Economics and Policy, Tokyo, between April 1997 and March 1998. The author has benefited since then from exchanges with Richard Cookson, Atsuo Kishimoto, Hideo Komatsu, Shuzo Nishimura, Richard Smith, Makoto Tamura, and Alan Williams. The usual disclaimer applies.

References

Blaug,M. (1992), *The Methodology of Economics: Or how Economists Explain* (2nd ed.), Cambridge University Press

Boadway,R. (1974), The welfare foundations of cost-benefit analysis, *Economic Journal*, 84:926-39

Boadway,R. (1979), The welfare foundations of cost-benefit analysis - A reply, *Economic Journal*, 86:359-61

Boadway,R., Bruce,N. (1984), *Welfare Economics*, Basil Blackwell

Cookson,R. (1996), Welfare economic dogmas, *Environmental Values*, 5: 59-74

Culyer,A.J. (1989), The normative economics of health care finance and provision, *Oxford Review of Economic Policy*, 5(1):34-58

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

Donaldson,C., Birch,S., Gafni,A. (1998), The 'distribution problem' in economic evaluation: Income considerations in the valuatoin of costs and consequences of health care programmes, paper presented at HESG meeting, Galway

Drummond,M,F., O'Brien,B., Stoddart,G.L.; Torrance,G.W. (1997), *Methods for the Economic Evaluations of Health Care Programmes,* 2nd ed., Oxford University Press

Foster,E. (1979), The welfare foundations of cost-benefit analysis - A comment, *Economic Journal*, 86:353-358

Gold,M.R., Siegel,J.E., Russell,L.B., Weinstein,M.C. (1996), *Cost-Effectiveness in Health and Medicine,* Oxford University Press

Johannesson,M. (1996), The relationship between cost-effectiveness analysis and cost-benefit analysis, *Social Science and Medicine*, 41(4):483-489

Johannesson,M. (1996), *Theory and Methods of Economic Evaluation of Health Care*, Kluwer Academic Publishers

Johansson,P-O. (1995), *Evaluating Health Risks: An Economic Approach*, Cambridge University Press

Krabbe,P.F., Essink-Bot,M.L., Gouke,J.B. (1997), The comparability and reliability of five health-state valuation methods, *Social Science and Medicine*, 45(11):1641-1652

Layard,R., Walters,A.A. (1994), Income distribution: Allowing for income distribution, in Layard,R., Glaister,S. eds., *Cost-Benefit Analysis* (2nd ed.), Cambridge University Press (Reprinted from Layard.R., Walters,A.A.,1978, *Microeconomic Theory*, McGraw-Hill)

Murray,C.J.L. (1996), Rethinking DALYs, in Murray, Lopez,A.D. eds., *The Global Burden of Disease,* Harvard University Press

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

Nord,E., Richardson,J., Macarounas-Kirchmann,K. (1993), Social evaluation of health care versus personal evaluation of health states, *International Journal of Technology Assessment in Health Care,* 9(4):463-478

Oka,T. (1997), *Welfare Economics and Environmental Policy* (*厚生経済学と環境政策*), Iwanami Shoten, in Japanese

Okuno,M., Suzumura,K. (1988), *Microeconomics II* (*ミクロ経済学II*), Iwanami Shoten, in Japanese

Olsen,J.A. (1997), Aiding priority setting in health care: is there a role for the contingent valuation method?, *Health Economics*, 6:603:612

Olsen,J.A., Smith,R. (1998), *Who* have been asked to value *what*?: A review of 54 WTP-surveys on health and health care, paper presented at HESG meeting, Galway

Patrick,D., Erickson,P. (1993), *Health Status and Health Policy: Quality of Life in Health Care Evaluation and Resource Allocation*, Oxford University Press

Pauly,M.V. (1996), Valuing health care benefits in monetary terms, in Sloan,F.A. ed., *Valuing Health Care: Costs, Benefits and Effectiveness of Pharmaceuticals and Other Medical Technologies*, Cambridge University Press

Richardson,J., Nord,E. (1997), The importance of perspective in the measurement of quality-adjusted life years, *Medical Decision Making*, 17:33-41

Sugden,R., Williams,A. (1978), *The Principles of Practical Cost-Benefit Analysis*, Oxford University Press

Wagstaff,A.(1994), QALYs and the equity-efficiency trade-off, in Layard,R., Glaister,S. eds., *Cost-Benefit Analysis* (2nd ed.), Cambridge University Press (Reprinted from *Journal of Health Economics*, 10: 21-41, 1991)

Weinstein,M.C. (1995), From cost-effectiveness ratios to resource allocation: Where to draw the line?, in Sloan,F.A. ed., *Valuing Health Care: Costs, Benefits and Effectiveness of Pharmaceuticals and Other Medical Technologies*, Cambridge University Press

Williams,A.(1972), Cost-benefit analysis: Bastard science? And/or insidious poison in the body politick?, *Journal of Public Economics*, 1:199-225

Williams,A. (1988), Ethics and efficiency in the provision of health care, in Bell,J.M., Mendus,S. eds., *Philosophy and Medical Welfare*, Cambridge University Press

Williams,A. (1997), Intergenerational equity: An exploration of the 'fair innings' argument, *Health Economics,* 6:117-132

Tables and Figures

Figure 1: Monetary evaluation of a given health improvement

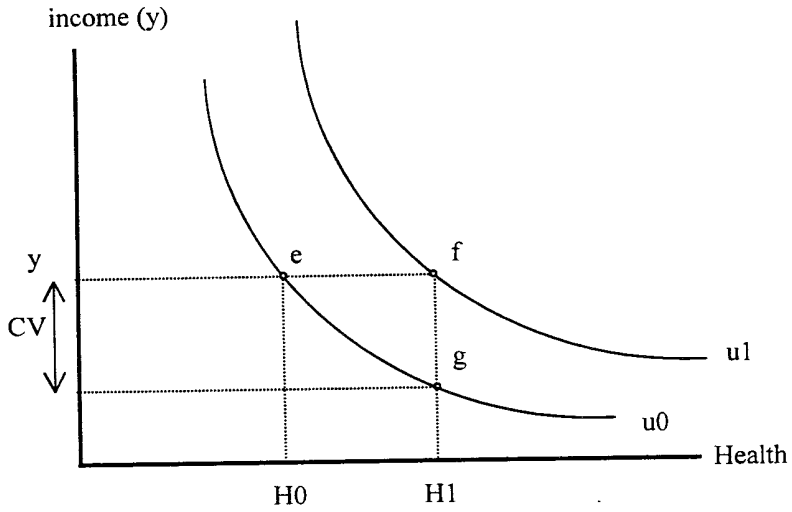

Figure 2: Potential for compensation after change

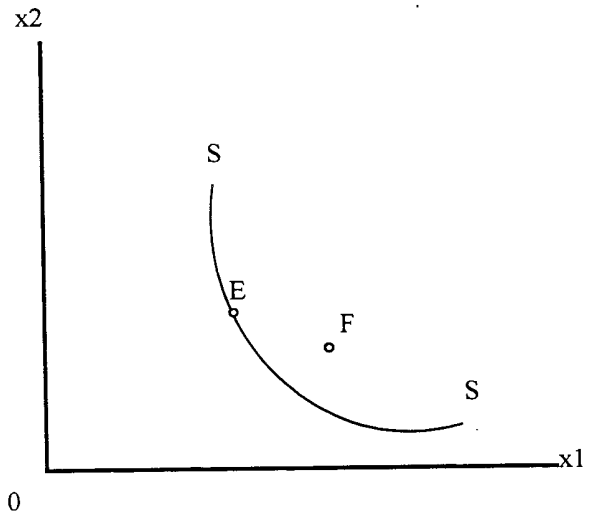

Figure 3: ΣCV>0

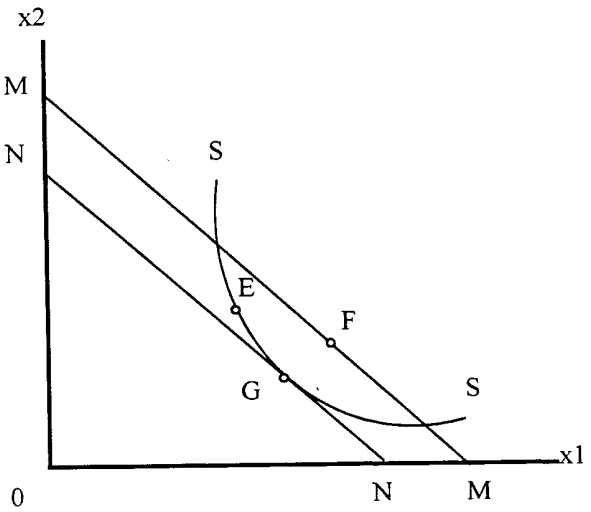

Figure 4: Weak and strong PCC

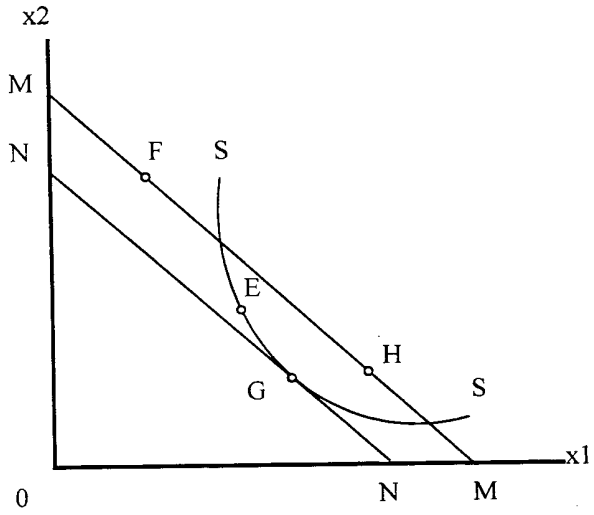

Figure 5: Constant but different marginal utility across individuals