How to value a person's life

John Broome

The 2022 Brocher Lecture

Introduction

Economists began making calculations involving the value of people's lives in the 1960s and 1970s. There have always been two strands of thinking. On the one hand there were transport economists, environmental economists and others who started incorporating the value of lives into cost-benefit analysis on the basis of people's willingness to pay to reduce their risk of dying.[i] On the other hand there were health economists who developed measures of the benefits of health care– principally qalys – to use in cost-effectiveness analysis of different treatments.[ii]

Recently there have been some signs of convergence between the strands, and I hope to make a small contribution to reconciling them. This is particularly worth doing because the economic value of life has acquired much greater importance in recent decades. For one thing, it is a major component of the social cost of carbon, which is the key parameter in climate change economics. And climate change is the leading problem of our age.

Willingness to pay

The first strand of thinking is embodied in the notion of 'the value of a statistical life'. Like many other people, I hate that term. But the part of it I hate is different from the part many other people hate. Many of them hate the word 'life', because they don't like to be seen as setting a value on people's lives. They prefer to set a value only on a *risk* of losing one's life. Their implication seems to be that preserving statistical lives is not as valuable as preserving real lives.

But I hate the word ‘statistical’. It reminds me irresistibly of the well-known remark attributed to Joseph Stalin:

A single death is a tragedy; a million deaths is a statistic.

Contrary to what Stalin implies, a million deaths is a million tragedies. The badness of deaths is proportional to the numbers of deaths.

Correspondingly, the badness of a risk to life is proportional to the probability of dying. The badness of a 1 in 10,000 risk of losing one’s life is just 1/10,000 of the badness of losing one’s life. This is an elementary consequence of expected utility theory. Why is a risk bad? Because of the badness of whatever it is a risk of. The primary object of value is what may happen – the possible outcome of the risk. The value of the risk derives from the value of this outcome. More, precisely, the badness (negative value) of a risk is the badness of the whatever it is a risk of, multiplied by its probability. That is to say, the badness of a risk of death is proportional to the probability of death.

For instance, an implication of expected utility theory is that the badness of exposing each of 10,000 similar people to a 1/10,000 risk of dying is the same as the badness of exposing on person to a certainty of death. I shall later qualify this conclusion on grounds of fairness. But for the time being I shall suspend the qualification and stick with this basic conclusion of expected utility theory.

It conflicts with traditional cost-benefit analysis. Traditionally, cost-benefit analysis values a benefit to a person by how much money the person would be willing to pay for it, and it values a harm to her by how much she would accept as compensation for bearing it. These amounts are technically the compensating variations of the benefit or harm. Traditionally, cost-benefit analysis reckons a change

that brings benefit to some and costs to others, as a good thing if and only if the sum of all the people CVs is positive. The sum of CVs is the criterion for identifying whether or not a change is a change is an improvement.

Now, the CV of a risk of death is not proportional to the probability of death. Valuation by the CV is therefore not consistent with expected utility theory.

It's easy to see why the CV is not proportional to probability. It is an implication of expected utility theory itself, applied to the person's own decision making. You can do the algebra, but the reason is easy to see without the algebra. Imagine you have to compensate someone for bearing a risk of death, and think how much the compensation she would require will increase as the risk gets higher. If she dies, she will get much less benefit from the compensation than if she lives, because she won't get to spend it. Indeed money may be worthless to her if she dies. As the risk gets higher, the expected benefit she receives from any particular amount of compensation therefore gets less and less, because the chance increases that she won't get to spend it. So she will require proportionally more compensation to make up for the chance of getting less benefit. In the extreme, if money is totally worthless to her if she dies, it will be completely impossible to compensate her with money for a very high risk of dying.

To put it briefly, the value of money to the person diminishes as the probability of her dying increases. It gets progressively harder to compensate her because you are you are trying to do so using a medium that has progressively less value to her. This makes it obvious that the CV is not a good measure of the value of risk. If you value risk of dying using the 'measuring-rod of money' as A. C. Pigou calls it, you will

find your measuring-rod constantly varying in length. It cannot work properly. The value of risk is proportional to the probability. Measuring value by the CV implies it is not, but this only shows the CV is an unsatisfactory means of measurement.

Oddly, the CV method of measurement is historically what led economists to concentrate their attention on the value of a *statistical* life. Lots of projects cause deaths. For example, big engineering projects very often lead to deaths in the course of their construction. These deaths are a cost of the project. If you value them by their CVs, you get very big costs because it takes a very large – probably infinite – amount of money to compensate someone for dying. So if you think the CV is a correct measure of value, you will think that any project is bad if it causes a death. But that is clearly not so. Some projects that cause deaths are worthwhile nonetheless. So what do you do?

What you ought to do is realize that the CV is not a good measure of value. But that is not what economists did, because traditional cost-benefit analysis is backed by a strong ideology. Instead they decided to measure the value of the risk of death rather than the value of deaths themselves. This cut out those very high valuations and allows some projects to count as beneficial even if they cause deaths.

This move to valuing statistical rather than individual lives was made in order to preserve the ideology. But it wasn't worth preserving. Traditional cost-benefit analysis based on the sum of CVs should have been abandoned long ago, for many reasons. I've said it is inconsistent with expected utility theory. This is a bad fault, but it pales into insignificance compared with some of its other faults. As long ago as 1941, Tibor Scitovsky showed that it leads to flatly contradictory

results.[iii] The sum of CVs in moving from some situation B to another A may be positive, and at the same time so may the sum of CVs in moving from A to B. The sum of CVs criterion therefore implies that A is better than B and also that B is better than A. This is a reductio ad absurdum of the criterion.

This particular absurdity can be circumvented by applying a double – backward and forward – criterion. One option A is declared better than another B if the sum of CVs in moving from B to A is positive and the sum of CVs in moving from A to B is not positive. But in 1955, Terence Gorman showed that this double test can imply a different sort of inconsistency.[iv] It can imply that A is better than B, B better than C and C better than A. So this double criterion is also shown to be false by reductio ad absurdum.

That was almost 70 years ago and it should have finished off traditional cost-benefit analysis. In any case, there was never a good argument in the first place for using the sum of CVs as a criterion. The sum of CVs is generally taken to be a test of whether the gainers from a project could compensate the losers, so that nobody ends up worse off. For this reason it is often called 'the compensation test' or the test of a 'potential Pareto improvement'. But actually it is not. Even if the sum of CVs is positive, the gainers may not be able to compensate the losers without someone's ending up worse off. This may surprise you at first, but remember that a transfer from the gainers to the losers changes the distribution of wealth and consequently it changes market prices. The sum of CVs is calculated at the prices that prevail before the transfer. If the gainers tried to make the transfer, prices would change and it might not be possible to end up with a Pareto improvement. This was demonstrated by Robin Boadway in 1974.[v] So the sum of CVs is not the compensation test.

Besides, there was never any good reason for accepting the compensation test in the first place. It was recommended by Nicholas Kaldor[vi] and supported by John Hicks[vii] in 1939, but neither of them offered any real argument for it. It is easy to produce counterexamples in which the gainers from a change could compensate the losers, but nevertheless the change is obviously not for the better.

In sum, the basis for traditional cost-benefit analysis – using the sum of CVs as a criterion – was thoroughly discredited decades ago. Traditional cost-benefit analysis should be abolished. The sum of CVs had ideological support because it purports to value a project without the need for interpersonal comparisons of wellbeing, even when the project is good for some people and bad for others. That ambition has to be abandoned. It was hopeless from the start. When a project is good for some and bad for others, obviously we have to compare the good of some with the bad of others. That is exactly what we are doing when we evaluate the project.

This does not imply that the CV of a risk to life is useless for the purpose of valuation. It provides useful information about the value a person sets on her life. But the value of a project cannot be found simply by adding CVs.

If we are to use CVs (willingnesses to pay) for valuation, we have to be ready to make adjustments according to the value of money to different people. It is widely recognized that the value of money to rich people is less than to poor people because the rich already have a lot of the things money can buy. Also, I have just explained another source of variation in the value of money. It depends on how near death a person is: if she is old or for some other reason exposed to a bigger risk of dying soon, money is worth less to her because she is less likely to have time to spend it. People's CVs must be adjusted

according to these differences in the value of money.

For instance, it is obvious the CV of reducing risk will on average be lower in a poor country than a rich one. For this reason, in an international project, saving life in a poor country will get lower priority than saving life in a rich one if we apply the criterion of the sum of CVs. But it's equally obvious that the reason the CV is lower in a poor country is that on average money is more valuable to the people there. So if we apply an appropriate adjustment, the same priority will not emerge.

Fairness

Now back to a point I made earlier and immediately suppressed. I said it was an implication of expected utility theory that the badness of exposing 10,000 people to a 1/10,000 risk of dying is the same as the badness of exposing one person to a certainty of death. There is reason to think on grounds of fairness that expected utility theory goes wrong here. It is better for the risk of death to be widely distributed rather than focussed on one person because it is fairer. That is the argument Peter Diamond made against expected utility theory in 1967,[viii] and it is a good one.

It even supplies a sort of backhanded support to using the sum of CVs. Just because the unadjusted CV of risk increases more than in proportion to the risk, it will reckon a more concentrated risk as worse than a more widely distributed one. Indeed, in 1982 the UK National Radiological Protection Board reached exactly this conclusion by this method.[ix] It was assessing the badness of radiation leaking from nuclear plants, using valuation by CV. If there were to be a particular number of deaths, it much preferred them to be widely distributed across the UK population, rather than concentrated on the

close neighbours of the plant. This conclusion aligns with what fairness might also recommend.

But this gives no real support to valuation by the sum of CVs. The CV of a risk has nothing to do with fairness. Fairness and the sum of CVs may sometimes imply the same conclusion, but that is just a coincidence. The NRPB's reason for preferring a wide distribution of deaths is specious.

Still, it is a real question what we should do about fairness. I think Diamond was wrong to see it as an objection to expected utility theory. I think the best practical way of taking fairness on board is to treat it separately it from goodness. Expected utility theory applies to good excluding fairness, and fairness has to be taken into account separately.

To take a much-discussed example: suppose a health service has life-saving treatment available, but not enough to treat everyone who needs it. If it is decided on grounds of qalys whom to give it to, it will go to people who are otherwise in good health rather than those who have other health problems, just because they have better lives. But it is unfair to deny life-saving treatment to some people just because their lives are already less good than they might be.

We should not allow this consideration of fairness to distort our judgement about the benefits of the treatment. It does indeed do more good to treat people who are otherwise healthy than those who are not. But this is not everything; fairness also matters. Sometimes we ought not to do the best thing, because doing so is too unfair. That seems to be so in this particular case: on grounds of fairness we should not discriminate against the less healthy candidates for treatment.

The upshot is that fairness does not constitute an objection to what I have said about the sum of CVs as a criterion in cost-benefit analysis. My argument was concerned with value

or goodness, but we have to remember that fairness is a further consideration that needs to be taken into account. I now revert to goodness, excluding fairness, for the rest of this lecture.

The value of a person's life

I come to the alternative historical strand of thinking about the value of life.

One feature of the alternative approach is that it is more interventionist. It involves thinking about how good lives actually are, recognizing that people may make mistakes about the goodness of their own lives. Economists like to base their valuations on people's preferences alone. But in thinking about the value of lives, we have to remember that people have different preferences at different stages of their lives. At the very least, those preferences will have to be integrated together to reach a value for the life, and this integration cannot itself be based on these preferences. So some intervention is inevitable.

A second feature of the alternative approach is that risk is not essential to the method of valuation. Of course, risk and uncertainty are always present, but they are taken into account in a more conventional way. When there is risk, various outcomes are possible. Each of these has a value and a probability, and together their values and probabilities determine the value of the risk in the way described by expected utility theory. The primary bearers of value are the possible outcomes, which are states of affairs that themselves have no risk in them. So the first job for this approach is to work out the value of life in a state of affairs that has no risk or uncertainty. Any practical decision will require risk to be accounted for later.

Let me illustrate the general problem of valuation as I see it. Even to illustrate it, I need to make an assumption. I shall assume that the value of the world depends only on each person's temporal wellbeing at each time she is alive, by which I mean how well the person's life is going at each time.

This assumption sets aside the wellbeing of animals, the intrinsic value of nature, and the value of human cultures in so far as they have values beyond their effects on human wellbeing. It even sets aside the value that the human species may have, apart from the wellbeing of the individuals who make it up. So it may not capture all the badness of human extinction. I am not denying the existence of all these values, but I am assuming that they can be taken into account separately from the value of human wellbeing. I've chosen to concentrate on the value of human wellbeing among other values.

I also mean to allow for different conceptions of wellbeing, from a hedonist conception consisting of pleasure less pain, to a very general conception that includes health, access to travel and social life, having as nice house, and so on. Given all these caveats, I think the assumption that the value of a state of affairs depends only on people's temporal wellbeings is fairly uncontentious.

If you grant it, I can illustrate the general form of our problem with a picture.[x] This picture is supposed to illustrate the problem of climate change. It shows two possible states of affairs. Time is measured horizontally, with the vertical line marking the present. Each horizontal line belongs to a person, and the graph sitting on that line shows the person's temporal wellbeing through her life. There are some presently-existing people and some future people. Some people exist in one possible world and not in the other. In the world of *business*

as usual, lives are better in the near future than in *respond* because responding to climate change requires a sacrifice, but the quality of life in the further future is less good, and lives are shorter.

Next I assume separability of people. That is, I shall assume that the goodness of the world is made up of the goodness of each of the people's lives. Again, this is not a very contentious assumption. It doesn't rule out causal interactions between people. Each person's temporal wellbeings may well be affected by how other people's lives are going; that is not excluded. But once we have identified the temporal wellbeings, we can evaluate the state of affairs person by person. This means we can split our task into two steps. First, work out how good is each person's life. Then work out how the goodnesses of all the people's lives go together to determine the goodness of the state of affairs.

The second step is the business of a social welfare function. For instance, we might adopt an egalitarian or prioritarian or utilitarian formula. But that's just the beginning; we then have to take account of changes in population. For example, some people adopt average utilitarianism as their social welfare function; others total utilitarianism. All this is very difficult, but fortunately I'm not concerned with the social welfare function. I am concerned with the first stage, to work out the goodness of the individual lives.

So what is the value of a person's life? I have already assumed implicitly that it depends on how the life goes, which is to say the person's temporal wellbeings at each time. I said that was fairly uncontentious, but the contention starts when we come to working out the form of the function from temporal wellbeings to the overall value of a person's life. How do temporal wellbeings aggregate? The simplest

function is just additive: the value of a life is the arithmetic total, or integral of its temporal wellbeings. We may call this 'intrapersonal utilitarianism'. But lots of other functions are possible, which take account of the shape of the life. For example, it may be better for life to get progressively better rather than progressively worse. It may be that the end of life is particularly important in determining how good the life is as a whole. Alternatively, the beginning may be the most important, and later times of life may be discounted. It may be good to have a high peak, or alternatively it may be good for a life to maintain an even tenor. And so on.

I have to confess that I know of few theoretical arguments that adjudicate among all these possibilities. Mostly it seems down to intuition to settle on the correct formula. I think there are good arguments for interpersonal utilitarianism, which is the view that social value is the sum of individual wellbeings. But similar arguments for intrapersonal utilitarianism are much less convincing.

Nevertheless, I suggest we adopt intrapersonal utilitarianism as a default theory, in the absence of an argument to show it's wrong. I don't insist it is correct. I only suggest we need some good reason for departing from it.

If it is right, then the value of extending a person's life is simply the total wellbeing she acquires during her extra period of life. Putting it another way, it's the total of wellbeing-adjusted life years in that period. I suggest this as a default because it seems the most conservative, neutral formula. It is also intuitively attractive. What could be more natural than to think the goodness of a life is its total goodness, integrated over time? Clearly many people working on the value of life have taken it for granted. For decades public health economists and others have taken for granted

more specialized versions of it, in the form of qalys and dalys. Although there has always been debate about the right way to make the quality adjustment in qalys, there's not been much disagreement about adding up years.

Interpersonal comparisons

So the value of a life can be described as the total of wellbeing-adjusted life-years. For practical decision-making, of course, the difficult bit is the wellbeing adjustment. But even before we get to that, there is an important practical implication of valuing lives on the basis of life-years. Even traditional cost-benefit analysis is equipped to work with life-years instead of undifferentiated life-saving. The notion of the VSLY – the value of a statistical life-year – is well recognized. It would surely be better to work with VSLYs rather than VSLs. Even if intrapersonal utilitarianism is not correct, saving a life is much less valuable if it extends the life by just a few weeks than if it extends it by many decades. This is a minimal improvement to practice that could easily be adopted. It should be. I know of some resistence to this practice, which is based on the fact that older people may well be willing to pay as much as younger people for reducing their risk of dying. But this is plainly because money is less valuable to them because they have less time to spend it, and they are also likely to have more money than the young.

Now, what about a practically implementable measure of temporal wellbeings? I am not going to give a definitive answer to this question. I could not give one in any case because it obviously depends on what a person's temporal wellbeing consists in. There are many theories about this, which have been much debated. Any answer to the question of measuring wellbeing has to be tied to a particular theory of

what wellbeing is.

To narrow the task, I shall look for a reconciliation with the first strand of thinking about the value of life. I want to stay as close as possible to the conventional methods of cost-benefit analysis, because this will make my suggestion easiest to implement in practice.

Conventional methods use the market prices of goods for valuation when they can. The reason for that is that prices measure wellbeing. The price a person pays for some good is supposed to measure the marginal contribution that good makes to her wellbeing. More exactly, the ratio between the price of one good and the price of another is equal to the ratio of the contribution the two goods make to each person's wellbeing. This is to assume that people's wellbeing is aligned with the preferences that underlie her choices. That might be either because her wellbeing actually consists in the satisfaction of her preferences, or that her preferences are accurately formed on the basis of her wellbeing. If we are to be reconciled with conventional cost-benefit analysis, we shall have to accept this basic assumption.

Sticking to conventional methods puts a demand on our measure of the value of life. It has to be commensurate with the prices of the ordinary goods that figure in the costs and benefits that are included in cost-benefit analysis. Our measure of temporal wellbeing will have to be such that the prices of goods are proportional to their marginal contribution to wellbeing as we measure it.

To see the point of this requirement, notice that money-metric utility or equivalent income as a measure of wellbeing does not meet it. Money-metric utility measures wellbeing as a quantity of money,[xi] which is a good start, but the prices of ordinary goods are not proportional to their

contribution to wellbeing measured this way. It's the reference prices embedded in the equivalence measure that are proportional.

Willingness to pay is a way to satisfy this requirement of commensurateness with prices. After all, prices themselves are willingnesses to pay. The price of a good is what people are willing to pay for the good at the margin. We can build a measure of a person's temporal wellbeing on the basis of her willingness to pay to extend her life at that level of wellbeing. That is my suggestion.

To carry it out, we shall need a person's willingnesses to pay to extend her life, for various types of life the person might lead during the extended period. By a 'type of life', I mean those natural features of a period of life that contribute to determining the person's temporal wellbeing during the period. Each type of life is assigned value by means of the person's willingness to pay for a marginal extension to her life of this type. These willingnesses to pay must be discovered empirically. For instance, people might be asked what is their marginal willingness to pay to extend their life by a marginal period of life of such-and-such a type. (The extension need not be at the end. It might be inserted in the middle of a life.)

These willingnesses-to-pay will give a value to each type of life, in terms of money. These values will be measured on a ratio scale. The zero of the scale is given by the life's not being extended at all. The scale will be particular to each person, and it will assign a value to various types of life the person might possibly lead.

We need different people's scales to be comparable. So we next need to bring their different scales into line. The zero of the scale is assigned to life's coming to an end, which is equally bad for everyone. So this zero level is already

interpersonally comparable. Consequently, it is only the size of the unit of value that remains to be aligned between people.

I assume that leading a particular type of life is equally as good for one person as it would be for anyone else. If there is a type of life that is possible for everyone, this type will have a place in everyone's scale of value. We have only to adjust each person's scale to make sure this type gets the same value for everyone, and then we shall have fully comparable scales.

If there is no type that is possible for everyone, the interpersonally comparable scale will have to be built by a sequence of pairwise comparisons. Each person can have her scale aligned with another person who can live one life of the same type as she can. We can hope that the whole population can be covered by overlapping pairs like this. If so, we can achieve fully comparable scales this way.

Since health is a component of temporal wellbeing, an interpersonal scale of health is a useful prototype. The scale of health used in qalys is built on a similar assumption to the one I made. Let a 'health-type' of life be the set of features of a period of life that contribute to determining how healthy a person is. We assume that two people are equally healthy if their lives share the same health-type. That is how levels of health are made interpersonally comparable.

Aligning people's scales in the way I have described is a way of adjusting each person's money values – her willingnesses to pay – according to the value of money for the person. All of a person's money-values are adjusted, which means that all her relative values remain the same. The value of extending her life relative to other goods such as food is unaffected by the adjustment.

If we average in some way across the populations of each country, the adjustments I have described give us an exchange

rate between the countries' currencies. Let us call it the 'value parity' rate. The rupee to dollar value parity rate will be much higher than the rupee to dollar purchasing-power parity rate. Purchasing-power parity makes the rupee price of goods the same as their dollar price. But since people in India are poorer than people in the US, goods are more valuable to them. So purchasing-power parity undervalues Indian people's goods, including their lives. But at the value parity rate, equally good lives will be accorded equal value.

An extremely crude, simplified version of this proposal restricts types of life to health-types, and ignores other features of lives. It assigns the same value to a qaly for each person. Between countries, exchange rates will be set to make this so. The qaly serves as a numeraire, that is to say. This is plainly a very rough approximation, since not all qalys do actually have equal value. Some people living in a particular health state are a lot better off than others in that same health state. However, it probably approximates the truth much better than assuming all dollars have equal value, which is the assumption implicit in traditional cost-benefit analysis.

So even the crudest, simplest version of my suggestion will lead to cost-benefit analysis that is better in this important respect than the traditional method.

References

Boadway, Robin, 'The welfare foundations of cost-benefit analysis', *Economic Journal*, 84 (1974), pp. 926–39.

Broome, John, *Weighing Lives*, Oxford University Press, 2004.

Diamond, P. A., 'Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility: comment', *Journal of Political Economy*, 75 (1967), pp. 765–6.

Drèze, Jacques, (1962) 'L'utilité sociale d'une vie humaine', *Revue Française de Recherche Opérationelle*, 23 (1962) pp. 93–118.

Fleishman, A. B., and M. J. Clark, 'Evaluating future detriment from radioactive discharges: judgements and implications for optimisation of protection' National Radiological Protection Board Report NRPB–R132 (1982).

Fleurbaey, Marc, 'Equivalent income', in *The Oxford Handbook of Well-Being and Public Policy*, Oxford University Press, 2016, pp. 453–75.

Gorman, W. M., 'The intransitivity of certain criteria used in welfare economics', *Oxford Economic Papers*, 7 (1955), pp. 25–35.

Hicks, J. R., 'The foundations of welfare economics', *Economic Journal*, 49 (1939), pp. 697–712.

Kaldor, Nicholas, 'Welfare propositions of economics and inter-personal comparisons of utility', *Economic Journal*, 49 (1939), pp. 549–52.

MacKillop, Eleanor, and Sally Sheard, 'Quantifying life: Understanding the history of Quality-Adjusted Life-Years (QALYs)', *Social Science and Medicine*, 211 (2018), pp. 359–66.

Scitovszky, Tibor de, 'A note on welfare propositions in economics', *Review of Economic Studies*, 9 (1941), pp. 77–88.

Notes

---

i. The first example of this strand that I know is 'L'utilité sociale d'une vie humaine' by Jacques Drèze.

ii. The history of the development of qalys is described in

---

detail by Eleanor MacKillop and Sally Sheard in ‘Quantifying life’.
iii. ‘A note on welfare propositions in economics’.
iv. ‘The intransitivity of certain criteria’.
v. ‘The welfare foundations of cost-benefit analysis’.
vi. ‘Welfare propositions of economics’.
vii. ‘The foundations of welfare economics’.
viii. ‘Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility’.
ix. Fleishman and Clark, ‘Evaluating future detriment from radioactive discharges’.
x. It is taken from my *Weighing Lives*, p. 10.
xi. See Fleurbaey, ‘Equivalent income’.