

## *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 analyses on the basis of people's willingness to pay for extending their lives.<sup>1</sup> On the other hand there were health economists who developed measures – principally *qalys* – for the benefits of health care to use in cost-effectiveness analysis of different treatments.<sup>2</sup>

Recently there have been some signs of convergence between the strands, and I hope to make a small contribution to their reconciliation. 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 figure 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, rather than on losing a life itself.

But I hate the word 'statistical'. It reminds me irresistibly of Joseph Stalin's famous (apocryphal) remark:

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.

And that is true of risks too. 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 the outcome. The nature of the derivation is easy to state: the badness (negative value) of a risk is the badness of the bad thing 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, it is an implication of expected utility theory that the badness of 10,000 people's being exposed to a 1/10,000 risk of dying is the same as the badness of one person's being exposed to a certainty of death (if all the people are similar). 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, which 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 accepting a project.

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. 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 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. So it will not work properly. The value of risk is proportional to the probability. Measuring value by the CV implies it is not, but this only shows it is an unsatisfactory means of measurement.

Oddly, though, this method of measurement is historically what led economists to concentrate their attention on the value of a *statistical* life. It happened around 1970. 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 numbers 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 means of valuation, you will think you have to reject any project that 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 there is a strong ideology behind traditional cost-benefit analysis, so that is not what economists did. 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 be accepted even if they cause deaths.

This move to valuing statistical rather than individual lives was made to preserve the ideology. 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.<sup>3</sup> 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.<sup>4</sup> It can imply that A is better than B, B better than C and C better than A. This criterion too is 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 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. 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.<sup>5</sup> 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<sup>6</sup> and supported by John Hicks<sup>7</sup> 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 purported 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. It does not *determine* the value of her life, but it can be good evidence about the value of her life.

If we are to use CVs (willingnesses to pay) as evidence, 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.

For instance, it is obvious that 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 10,000 people’s being exposed to a 1/10,000 risk

of dying is the same as the badness of one person's being exposed to a certainty of death. There is an argument to say that expected utility theory goes wrong here, because it is better for the risk of death to be widely distributed rather than focussed on one person. This is on ground of fairness. That is exactly the argument Peter Diamond used against expected utility theory in 1967,<sup>8</sup> 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.<sup>9</sup> 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 support for unadjusted CV valuation is ineffective. The CV of a risk has nothing to do with fairness. Fairness and the sum of CVs may reach the same conclusion, but that is just a coincidence. The NRPB's reason for preferring a wide distribution of deaths is specious.

Still, what we should do about fairness is a real question. 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 (or value). Of course, fairness is good, but it behaves in a very different way from other goods, so it is advantageous to separate it. Expected utility theory applies to good, 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. But it is unfair to deny life-saving treatment to some people just because their lives are already less good than they might be. However, 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 good 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 I described: on grounds of fairness we should not discriminate against the less healthy candidates for treatment.

The upshot of this is that fairness does not constitute an objection to what I have said about the sum of CVs as a criterion for 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 once more revert to goodness.

### **The value of a person's life**

Now I come to the alternative strand of thinking. Here I shall often speak in my own voice, since at a general level I subscribe to this strand.

Here are some features of this alternative approach. First, 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, and this integration cannot itself be based on the 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 always has to be taken into account, but it is 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 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 without risk or uncertainty. Any practical decisions 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 are separable from the value of human wellbeing, so they can be taken account of separately. I've chosen to concentrate on the value of human wellbeing.

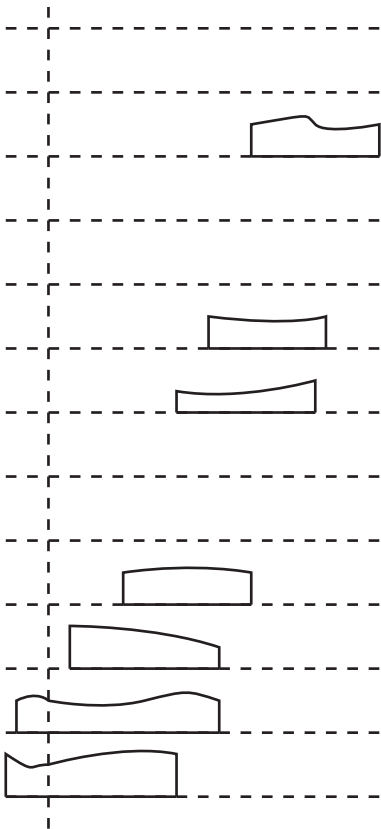
I also mean to allow for different conceptions of wellbeing, from a hedonist conception consisting of pleasure less pain, to a very general conception that includes health, access to travel and social life, having a nice house, and so on. Given all these caveats, I think the assumption that the value 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.<sup>10</sup> 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*, the quality of life in the future is less good than in *respond*, and lives are shorter.

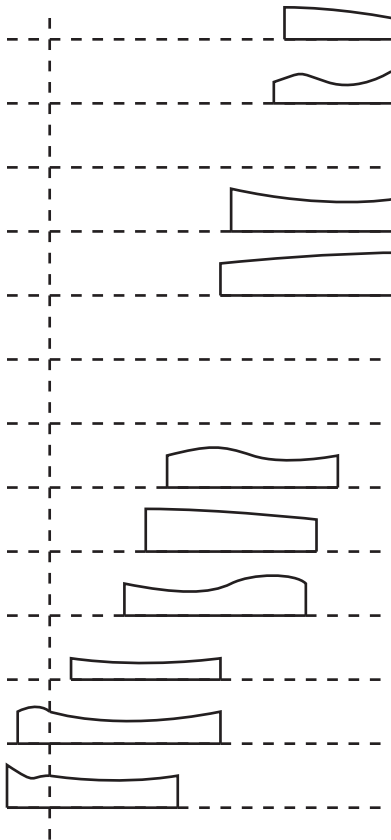
Next I assume separability of people. That is, I shall assume that the goodness of the world is made up of the goodness of each of the people's lives. Again, this is not a very contentious assumption. It doesn't rule out causal interactions between people. Each person's temporal wellbeings may well be 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 stages. 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.

This second stage is the business of a social welfare function. For instance, we might want an egalitarian formula, or one that does not favour equality, such as a utilitarian formula. But that's just the beginning; we then have to take account of changes in population. Some people adopt average utilitarianism, for example; others total utilitarianism, in which case they have to settle on a zero of wellbeing. 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



A: business as usual



B: respond

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 to a person's life. How do temporal wellbeings aggregate? The simplest function is just additive: the value of a life is the arithmetic total 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 how life ends 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 less so. 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 do think there are good arguments for interpersonal utilitarianism, which means 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 if we are to depart 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: waly as I playfully called them in my book *Weighing Lives*. 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 been little disagreement about the use of 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, of course. 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. Some practitioners seem nevertheless reluctant to adopt it. Perhaps this is because it favours saving the lives of young people over the lives of old people, and empirically the old may be just as willing as the young to pay to reduce risk to their lives. But if that is so, it is because old people have more money and less to do with their money. It is because money has less value to them, not because their lives have the same value to them as young people's lives have to them.

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

plainly 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.

An underlying principle of conventional methods is that prices and compensating variations in general can be understood as measures of wellbeing. The price a person is willing to pay for some good is supposed to measure the marginal contribution that good makes to her wellbeing. More exactly, the prices of goods are proportional to the relative contributions the goods make to each person's wellbeing. This assumes that the person's wellbeing is aligned with the preferences that underlie her choices. This 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 assumption.

Sticking to conventional methods consequently 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 included in cost-benefit analysis. Our measure of temporal wellbeing will have to be commensurate with ordinary goods, in such a way that the prices of goods are proportional to their marginal contribution to wellbeing.

To see the point of this requirement, notice that the equivalent income measure of wellbeing does not meet it. Equivalent income measures wellbeing as a quantity of money, which is a good start, but the prices of ordinary goods are not proportional to their contribution to wellbeing measured this way.

We can satisfy this requirement by treating a person's willingness to pay to extend her life as a measure of the temporal wellbeing she will enjoy during the period for which the life is extended. That is my suggestion.

It will be an empirical exercise to determine on the basis of available data what is a person's willingness to pay to extend her life. The data will be of the sort already used to calculate the value of statistical life years, including particularly the person's willingness to pay to diminish her risk of death. But we shall need much richer data because we need values for each type of life extension. We need willingnesses to pay for each different type of life the person might lead during the extended period. By a 'type of life', I mean the set of all those natural features of a period of life that contribute to determining the goodness of the person's life – which is to say her temporal wellbeing – during the period. Each type of life is assigned value by means of the person's willingness to pay for an extension to her life of this type.

In principle, the person's willingnesses to pay for extended life of different types will assign values to each of these types, 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 and instead coming to an end. This ratio scale is particular to the person, and it will assign a value to each type of life the person might lead.

We need different people's scales to be comparable. So we next need to bring different people's 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 needs to be aligned between people.

I assume that a particular type of life is equally as good for one person, if she leads a life of this type, as it is for anyone else who leads a life of that type. 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 a 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, the interpersonal scale of health is a useful prototype. The scale of health used in qalys is built on a similar assumption. 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. This makes levels of health 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, with means that all her relative values remain the same. For example, the value of extending her life relative to other goods such as food remain the same.

If we average in some way across the populations of different countries, these adjustments give us an exchange rate between the countries' currencies. Let us call it the 'value parity' rate. The rupee/dollar value parity rate will be much higher than the rupee/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.

A very crude, simplified version of this proposal treats all types of life as having the same value. It assigns one particular value to every life year, the same for each person. Between countries, exchange rates will be set to make this so. The life year serves as a numeraire. This is plainly a very rough approximation, since not all life years have equal value. However, it is probably a lot better as an approximation than assuming all dollars have equal value, which is implicit in traditional cost-benefit analysis.

So even the crudest, simplest version of my suggestion will lead to better cost-benefit analysis 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*, p. 210.
- 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 Opérationnelle*, 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).
- 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

1. The first example of this strand that I know is 'L'utilité sociale d'une vie humaine' by Jacques Drèze.
2. The history of the development of qalys is described in detail by Eleanor MacKillop and Sally Sheard in 'Quantifying life'.
3. 'A note on welfare propositions in economics'.
4. 'The intransitivity of certain criteria'.
5. 'The welfare foundations of cost-benefit analysis'.
6. 'Welfare propositions of economics'.
7. 'The foundations of welfare economics'.
8. 'Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility'.
9. Fleishman and Clark, 'Evaluating future detriment from radioactive discharges'.
10. It is taken from my *Weighing Lives*, 2004, p. 210.