The value of living longer

John Broome

University of Oxford
john.broome@philosophy.ox.ac.uk

For: *Health, Ethics and Equity*, edited by Sudhir Anand, Fabienne Peter and Amartya Sen, Oxford University Press.

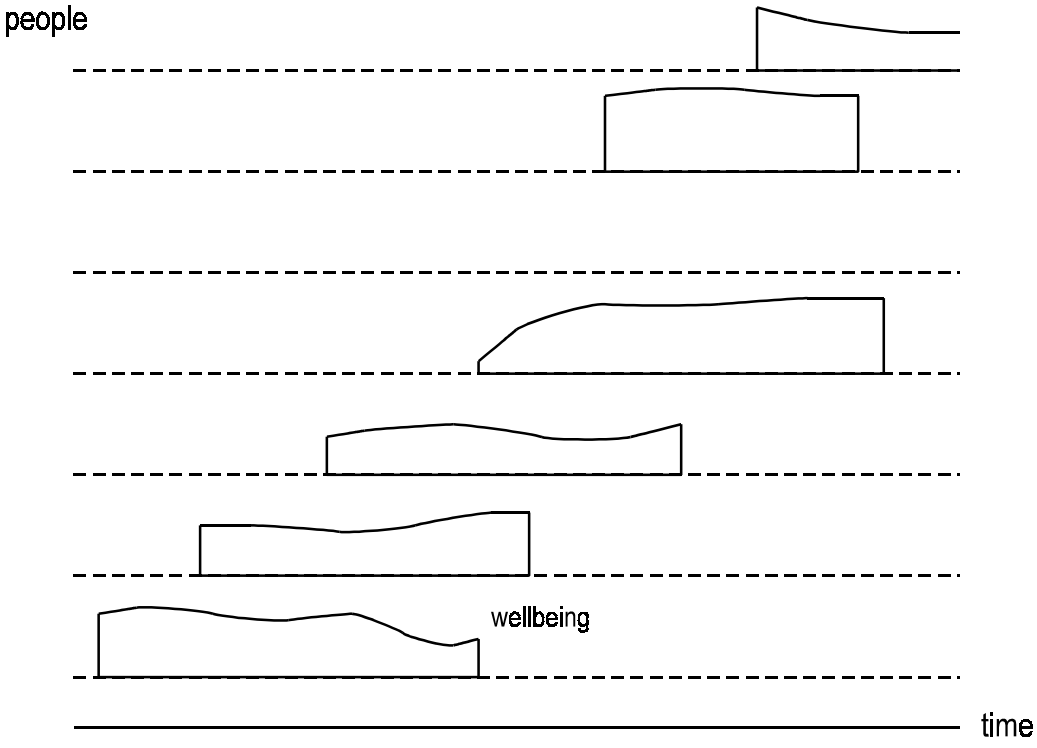
I have received extremely helpful comments from Francesca Silverton and the participants at the conference that led to this volume, particularly from Norman Daniels and from the editors. I am very grateful for them. Part of the work on this paper was done while I was a Visiting Fellow at the Swedish Collegium for Advanced Study in the Social Sciences. Another part was financed by the UK Economic and Social Research Council, under grant L32027301097. I thank both these institutions for their generosity.

1. Longevity and distributions of wellbeing

Longevity is an important aspect of health. Anyone who wants to measure the health of a population will need to take it into account. But there are special difficulties about doing so. This paper explains why.

Measuring health is one thing and valuing health is another. Which should we aim at? I assume we are interested in a person's health as a component of her wellbeing. Wellbeing consists of various different sorts of goods, such as health, comfortable living, freedom from oppression, and so on. So we are interested in health as one good thing to be set alongside others. We are therefore interested in how good it is, which means we are interested in its value. This paper is about valuing health – measuring the value of health – rather than measuring the quantity of health. Specifically, it is about the value of longevity.

Figure 1



To set a value on longevity is an aspect of a general problem in the theory of value: the problem of aggregating good or wellbeing.[1] Let us concentrate on a particular country. Each person who lives in this country is born at some time, dies at some time, and at each time during her life enjoys some degree of wellbeing. Figure 1 is a schematic illustration of the country's progress. Time is measured horizontally in the diagram. Stacked up vertically is a series of dotted horizontal lines. Each is the horizontal axis of a little graph, which shows the progress of a single person's wellbeing through her life. In these graphs, wellbeing is measured vertically. A person's graph starts at the time she is born, and ends at the time she dies. One of the dotted lines has no graph on it, for a reason that will appear later in this section. In sum, this diagram shows how wellbeing is distributed in the country, across people

and across time.

The problem of valuation, looked at in the most general way, is to set a value on distributions like this. We are trying to judge how well off is a country's population, and this must be determined by the wellbeing of each person at each time. So we need to *aggregate* wellbeing across a distribution; we need to take the wellbeing of each person at each time, and put all those quantities of wellbeing together to determine how good is the distribution as a whole.

In practice, rather than calculating the absolute goodness of a single distribution, we shall be more interested in making comparisons between distributions, to see which is better than which, and how much better. When things change, we need to know whether the change is beneficial or harmful, and the size of the benefit or harm. In this paper I am interested particularly in changes that increase or decrease the lengths of people's lives, to measure the overall harm or benefit that results.

Unfortunately, we cannot entirely isolate this single question of longevity. For one thing, in valuing longevity, we shall have to be guided by the principles that govern aggregation in general, and we need to keep in mind what those principles are. Also, there is one fact about longevity that we particularly cannot ignore. Changing the lengths of people's lives generally changes the number of people who are born. For example, as death rates fall and people live longer, they tend to have fewer children. Consequently, we cannot ignore changes in the country's population when we try to value longevity.

Figure 2

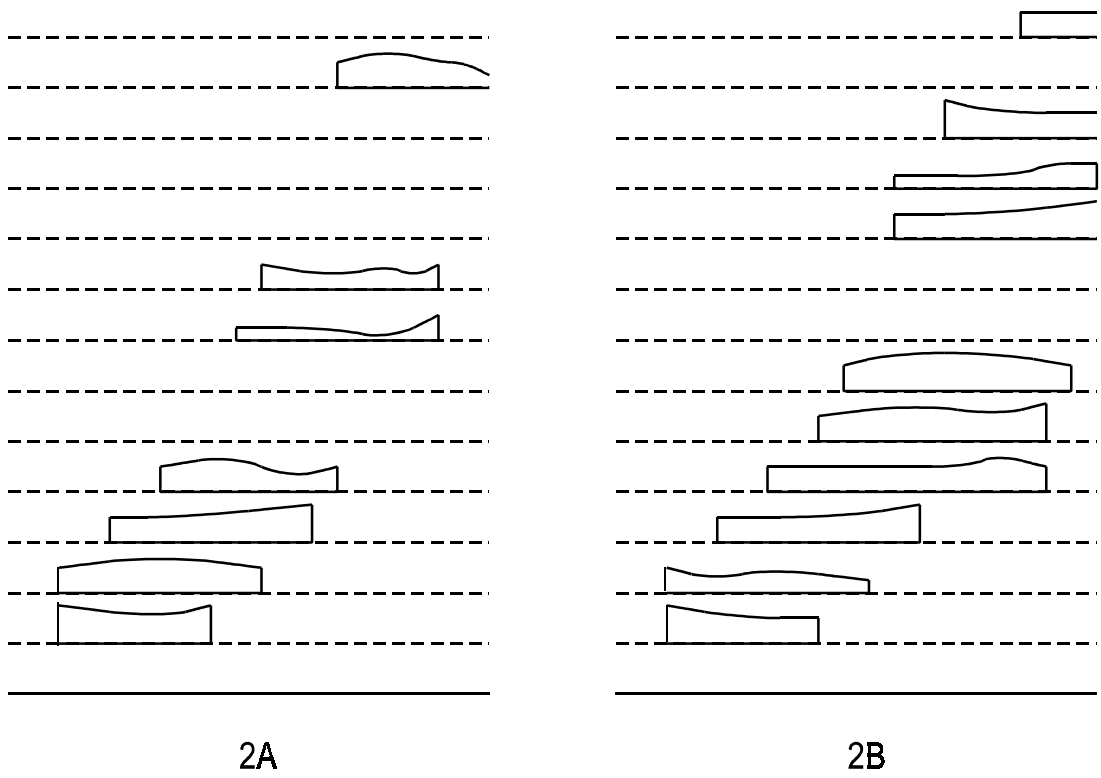


Figure 2 illustrates the general problem of comparison. It shows two alternative

distributions of wellbeing. People mostly live longer in one than in the other. One of the alternatives also has a larger population than the other, so some people exist in one who do not exist in the other. Each person is allocated a dotted line in the diagram; corresponding lines in the two halves of the diagram represent the same person. So when a person lives in one alternative and not the other, her line is left blank on one side of the diagram.

The problem of comparison is to determine which of two distributions like this is better overall. This paper is about those aspects of this problem that have to do with the length of people's lives.

2. Aggregation and separability

We need to take a whole distribution of wellbeing, and aggregate across time and across people, to arrive at an overall value. There are different ways we might set about doing that. We might hope to do it in two steps, aggregating across people as one step and across times as another. We might take these steps in either order. One approach is to start by aggregating across people, doing so for each time separately, and then to aggregate across times. In this approach, the first step values the state of the country at each time. The second takes all these values arrived at in the first step – one for each time – and aggregates them across times. An alternative approach is first to set a value on each person's overall wellbeing, by aggregating across times within the person's life. Then the second step is to take all these values arrived at in the first step, one for each person, and aggregate them across people.

The first of these approaches involves setting a value, at each time, on everybody's wellbeing taken together. Let us call this a 'snapshot' valuation, because it is made for a single moment in time, and let us call this first approach to valuing a distribution 'the snapshot approach'. To make a snapshot valuation, it is as though we photograph a diagram like figure 1 through a narrow vertical slit, which only reveals the distribution of wellbeing at a single time. Snapshot valuations are implicit in a great deal of our thinking about the progress of a country – for example, its progress in health. We talk about how a country's health improves from year to year, or perhaps how the inequality in its health increases. In doing this, we set a value on the state of the country at each time, and watch this value develop as time passes. So even if we are not intending to take the second step of aggregating the snapshot values across time, snapshot values commonly have a place in our thinking.

The snapshot approach to valuing a distribution only works properly if special conditions are met. Each snapshot valuation, made for a particular time, must make good sense. This means it must be possible to value the distribution of wellbeing in a country at a particular time, independently of people's wellbeing at other times. Furthermore, these snapshot valuations, one for each time, must together determine the overall value of the whole distribution: the sequence of snapshot valuations must contain all the information that is needed to determine the overall value. In technical language, this double condition is called *separability* of times.[2] The snapshot approach to valuation is feasible only if times are separable.

The appendix to this paper defines separability of times more precisely, and distinguishes it from another condition that is sometimes given the same name in economics.

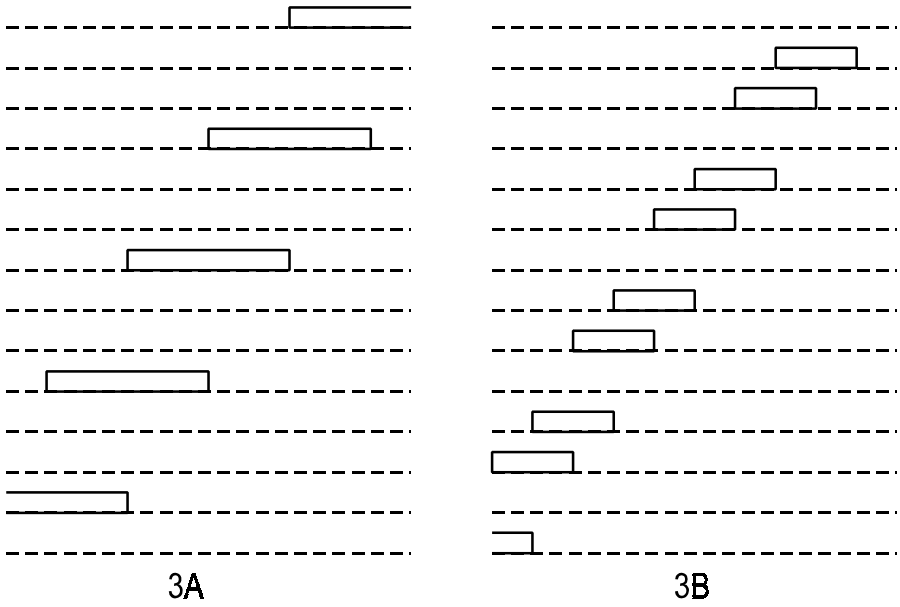
The alternative approach to aggregation also assumes a sort of separability: separability of people's lives. This approach starts by trying to find a value for each person's life by aggregating across times within her life. It is as though we look at the diagram through a horizontal slit. Then as a second step we aggregate across people. I shall call this the 'people approach' to aggregation. It will work properly only if a person's life can be valued independently of other people's wellbeing, and if the separate values for each person's life

together determine the overall value of the distribution. This double condition is called separability of lives.[3]

It is plausible that lives are separable. It is implied by something I call 'the principle of personal good'.[4] This principle is formally parallel to the well-known Pareto principle, but it is expressed in terms of what is good for people – people's wellbeing – rather than their preferences. The principle of personal good says that the value of a distribution of wellbeing depends only on the wellbeing of the individual people. To put it another way, overall goodness is a function of individuals' wellbeing. This implies we can value a distribution by first calculating each individual's wellbeing – aggregating across times in her life – and then calculating overall goodness on the basis of these amounts. So it implies separability of lives. Since the principle of personal good is very plausible,[5] separability of lives is plausible. This means the people approach to aggregation could be successful.

On the other hand, separability of times is not a plausible assumption, so the snapshot approach is dubious. At least, separability of times implicitly commits us to much more than most of us would be willing to accept. My principal aim in this paper is to make this point. It means we cannot legitimately make snapshot valuations of health.

Figure 3



3. The objection to separability of times

To see why times are not separable, think first about the example shown in figure 3. It shows two alternative distributions of wellbeing that are to be compared. In each, exactly two people are alive at each time. All the people who live in either alternative have the same constant level of wellbeing throughout their lives. The difference between the distributions is that each person in 3A lives longer than each person in 3B (twice as long, as I have constructed the

example).

The example is very stylized, in that the population is only two, and people live completely uniform lives. Still, it captures some of the features of a typical ‘demographic transition’. In a typical demographic transition, a country starts from a stable population with a high birth rate and a high death rate. Then the birth and death rates drop, and eventually the country arrives at another stable population, with lower rates. Figure 3 shows two possible states of a country, each with a stable population. The birth and death rates are high in the first and low in the second. So it is a stylized picture of a country before and after a demographic transition. One unrealistic aspect is that the population at any time is the same in both states: just two people. In practice, when a country passes through a demographic transition, its death rate falls before its birth rate. The result is that by the time its population reaches a new stable state, it is much larger. The example misses this feature, but it does represent another important feature of a transition: afterwards, people live longer.

Which of the two distributions is better? Let us try to compare them using the snapshot method. The method starts by comparing them time by time. We take each particular time on its own and without reference to any other time, and compare the values of the two distributions at that time. When we do this for any particular time, we see that the two distributions are identical at that time in terms of wellbeing. Each has two people alive, and each of those people has the same wellbeing. So we have to conclude that the two alternatives are equally good at that time. We must reach the same conclusion for every time. If times were separable, these snapshot comparisons would fully determine the comparative value of the two distributions. We would therefore be forced to conclude that the distributions, viewed as a whole, are equally good. That would be the consequence of separability of times.

We could not escape this conclusion by looking at the moments of birth and death. In my example, each person comes into existence at midnight; before midnight she does not exist and at midnight or later she does exist. Each person also goes out of existence at midnight; before midnight she exists and at midnight or later she does not exist. Consequently, there are no momentary overlaps when three people are alive, and no momentary gaps when only one person is alive.

So if times were separable, we would be forced to conclude that the two distributions in figure 3 are equally good. Yet that is surely wrong. Intuition suggests that longevity is valuable and consequently that 3A is better than 3B, because in 3A each person has a longer life.(Section 4 examines the basis of this intuition.) Therefore, times are not separable.

The example shows what is wrong with separability of times. For a person to live a longer life is a good thing. Yet the length of a person’s life is not something that shows up when we look at times separately. It is a feature of her life that only appears when we take a span of times together. So its value cannot show up in a snapshot valuation.

4. Separatism

The objection I made to separability is that 3A is better than 3B, which separability denies. But the basis of this claim was only the intuitive thought that longevity is valuable. One possible response to the objection is simply to reject it. This is a hard headed, intuitively implausible response, but it is possible. I shall call it ‘separatism’. Separatism is the view that times are separable.

When we take a snapshot of the wellbeing of the people in a country at a single time – surveying a vertical line in one of my diagrams – we see a lot of little pieces of wellbeing, each belonging to a different person. A snapshot does not allow us to see how these pieces are connected in a horizontal direction with other pieces; we do not see how pieces are packaged

together within individual lives. That is why the length of people's lives does not appear in a snapshot valuation. But the horizontal packaging of lives seems intuitively to make a real difference to value; longevity seems a good thing, for instance. The separatist denies this.

Separatism implies that horizontal connections make no difference. All that matters is the little pieces of wellbeing that appear at each time, not their packaging in lives. Consequently, longevity is not valuable. At least, it is valuable in one way but not another. Other things being equal, prolonging a person's life adds pieces of wellbeing to the distribution of wellbeing, and a separatist values those pieces. In that way, she values longevity. However, she gives no value to the fact that those pieces are joined on to a life that already exists, to make a longer life. Suppose a person were to die and a new person appear in her place and enjoy the same wellbeing. From the separatist point of view, that would be just as good as the first person's continuing to live would be. In *that* way, a separatist does not value longevity. She gives no more value to a long life than to several short lives that together make up the same period of time.

In figure 3, the same total amount of time is lived in 3A as in 3B. But in 3A it is apportioned to fewer longer lives, and in 3B to more shorter lives. A separatist values both options equally, because she does not care how the living is divided amongst lives.

The separatist view that horizontal connections make no difference is an extreme version of the view that personal identity does not matter, which was propounded by Derek Parfit in Part III of *Reasons and Persons*. Parfit arrives at his own less extreme view on metaphysical grounds. Some further philosophical arguments might bring us to the stronger conclusion that times are indeed separable.[6] I shall not review the arguments here. I only want to suggest that separatism could be given some philosophical basis, despite its implausibility.

The simplest example of a separatist theory of value may be called 'complete utilitarianism'. A complete utilitarian thinks that the value of a distribution is simply the total of all the wellbeing enjoyed at any time by anyone. To arrive at an overall value, we simply add wellbeing across the whole distribution – across people and across times. One distribution is better than another if and only if it has a greater total of wellbeing. A complete utilitarian does not care in any way about how wellbeing is distributed. For one thing, she does not care how it is packaged into individual lives. All that matters is wellbeing; who gets wellbeing is irrelevant.

Although separatism might receive a philosophical defence, I shall continue to reject it on grounds of its implausibility. From now on in this paper, I shall assume it is false.

5. Dispersing value

I have argued that times are not plausibly separable. But those who measure health will not give up snapshot valuations so easily. Apart from the tough separatist defence, there is a more pragmatic way of trying to preserve separability in the face of the objection I have raised.

I said the value of longevity does not show up when we look at times separately, but only when we take a span of times together. Consequently, it does not show up in snapshot valuations. However, one component of a person's wellbeing at any time is her health. Her health in turn consists of several components: functioning limbs, freedom from pain and so on. A person's longevity is one component of her health, so we might simply include it along with the other components at every time. We might say that someone who has a longer expectation of life is in this respect healthier at every time than someone who has a shorter expectation. So we might include her expectation of life within her health at each time, and hence within her wellbeing at each time. If we do that, when we make a snapshot valuation at any time, the person's longevity will show up in the snapshot. In this way, we might make

times separable, and snapshot valuations legitimate.

I want to keep the term ‘separatism’ for the hard-headed view that gives no value to longevity. So I do not count this new idea as a version of separatism, even though it supports separability. It is an example of a strategy within the theory of value that I call the ‘dispersion’ of value. For the sake of comparison, I shall mention a different application of dispersion, before assessing how successfully it copes with longevity.

The different application provides a defence of separability of lives, rather than separability of times. Separability of lives can seem to be threatened by the value of equality. I explained earlier that separability of lives is a consequence of the principle of personal good. According to this principle, the value of a distribution of wellbeing depends solely on the wellbeing of individual people. But many people think equality of wellbeing has value: a more equal distribution of wellbeing is better than a less equal one. One might think this value cannot belong to any individual person, because it depends on a relation between people: specifically on the difference between some people’s wellbeing and others’. It seems we first need to determine how well off each person is before we can see how equal the distribution is. Consequently, the value of equality must be a purely social value, separate from each individual’s wellbeing. So one might think, and the thought threatens separability of lives.

This argument about equality is investigated in Chapter 9 of my *Weighing Goods*. I argue there that if equality is indeed a good thing, it is so because it is good for individual people. Conversely, if inequality is a bad thing, it is so because it is bad for individual people. The badness of inequality is not a separate negative value beyond people’s own wellbeing. It is part of people’s wellbeing itself. This argument is an example of dispersion. At first, equality seems to be a separate, social value that cannot be captured within the wellbeing of individuals. But I took this apparently social value and dispersed it among the individuals. I argued it actually belongs to the individuals separately, despite first appearances.

An essential part of the strategy of dispersion is to justify it.[7] Take any value that apparently does not belong to people as individuals, so it seems to contradict separability of lives. Let it be biodiversity, say – suppose biodiversity is valuable. A theorist could always disperse this value in a formal way by arbitrarily dividing it into parts and allocating each part to a person. She could then say that each person’s allocated part was a component of that person’s wellbeing. But this would not be a sound move unless the theorist could demonstrate that biodiversity was indeed part of the person’s wellbeing. This demonstration needs to be made for the specific value that is in question. In the case of biodiversity, it might or might not be convincing. In the case of equality, I hope I justified dispersion successfully by producing an account of the value of equality in terms of fairness, which made it definitely an individual value.

Unless an instance of dispersion is shown to be genuine in this way, it will be no help with aggregation. If we know the aggregate value of a distribution, we can always disperse the value arbitrarily to people. But to do that, we would have to know the aggregate value first. Since we are trying to find aggregate value, this arbitrary manoeuvre would be pointless. On the other hand, if the value is genuinely a part of each individual’s wellbeing, we should be able to assess each person’s share of it independently, and then include these shares in an aggregation across people.

6. Dispersing the value of longevity

Now back to our original problem, and separability of times. On the face of it, the value of longevity does not appear at individual times. But we could attribute it to individual times by treating it as part of a person’s health at each time. That is to say, we take this value, and

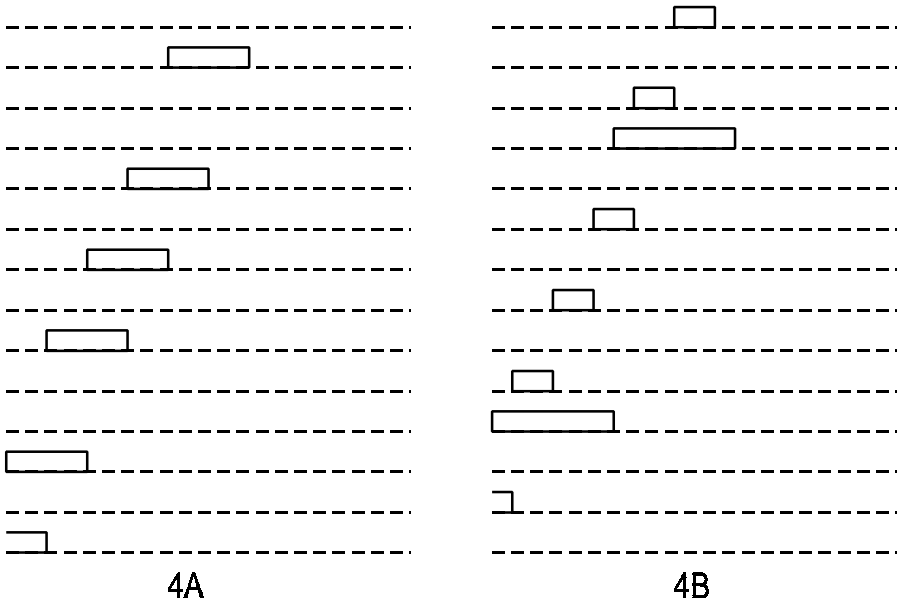
disperse it to individual times. In this way, perhaps we can make the value of longevity separable between times. That is the strategy.

If we knew the aggregate value of a distribution, including the value of longevity, we could arbitrarily divide it up and disperse it among the times separately. There are many arbitrary ways of dispersing the aggregate in such a way that the dispersed parts could be re-aggregated to reach the correct result. But as I say, an arbitrary dispersion is not enough, because it will not help with aggregation. The dispersion must be justified.

I doubt that dispersing the value of longevity across time could be justified. It seems to me that longevity is genuinely a value that does not appear at any particular time in a person's life; it is genuinely a feature of the life taken as a whole. But to give dispersion a chance, I shall waive this general scepticism about justification. The value of longevity might be divided up amongst times in various ways, and I shall examine some of them, without insisting that they be fully justified in advance. I shall ask only that they are not entirely arbitrary; they must make some sort of sense. I shall show that, quite apart from the matter of justification, they do not do what they need to do. They do not succeed in properly capturing the value of longevity.

This argument against dispersion will not be conclusive. I cannot rule out the possibility that someone might invent a convincing and successful way of dispersing the value of longevity. I can only say that the methods of dispersion I can think of fail. I shall also be able to explain why they fail, and this strengthens my argument. The explanation of why these particular methods fail can be generalized. It explains why dispersion in general can be expected to fail. So by the end of this section, we shall have good grounds for rejecting dispersion.

Figure 4



Take figure 4 as an example. It is a diagram of my usual sort, showing graphs of people's wellbeing. It shows wellbeing *without* taking account of life expectancy. That is to say, it does not include any dispersed value that we might attribute to longevity and add into the people's wellbeing at a time. The diagram shows two alternatives as usual. In both of them, two people are alive at each time, and their wellbeing apart from longevity is at the same constant level throughout their lives. In 4A, everyone lives for two years. In 4B, some people live for three years and others for one.

Distribution 4A is surely better than 4B. In 4A, everyone lives for two years. In 4B, three-quarters of the people live for only one year and the remaining quarter live for three years. In 4A, people's lifetimes are on average two years; in 4B, they are on average one and a half years. (The average lifetime of the people *who are alive at any particular time* is two years, but that is a different matter.) If we care about longevity, therefore, we will definitely favour the first alternative. This conclusion does not depend on caring specifically about the average. Look at it this way. In 4A, each period of twelve years lived is lived by six people. In 4B, each period of twelve years lived is lived by eight people. So the same period of life is divided amongst more people in 4B. If we value people's longevity in any way, we will certainly favour 4A.

There is plausibly a further consideration in favour of 4A: equality. People's lives are equally long in 4A but unequally long in 4B. If we value equality, that will give us reason to count 4A better. But it is a quite separate reason from the value of longevity. It stems from the value of equality in longevity rather than from the value of longevity itself. My example has the feature that the value of longevity is correlated with the value of equality in longevity. This is unfortunate, because it muddies the argument, but I cannot eliminate it. Unfortunately, it is inequality in lifetimes that causes dispersion to fail, but in a way that is independent of the value of equality itself; I shall explain why at the end of this section. So instead, I propose to isolate the value of longevity by ignoring any value that equality may have. For the sake of argument, then, let us assume equality has no value.

I say 4A is better than 4B, on grounds of longevity. Not everyone would agree. If you value only the total amount of wellbeing, independently of whom it comes to, or how it is packaged into lives, then you will think the two distributions in figure 4 are equally good. Both have the same amount of wellbeing altogether. Your view is then the one I called complete utilitarianism. It implies separatism, and it gives no value to longevity. If you take this view, you already believe times are separable, without dispersion. You think dispersion makes no contribution to separability, because separability is correct anyway. In this section, I am not addressing you. I am addressing people who think longevity is valuable, but nevertheless hope to preserve separability of times by means of dispersion.

These people will want to reproduce the conclusion that 4A is better than 4B. They are hoping to do so by means of snapshot valuations, taking each time separately and comparing the distributions at that time. They want to bring longevity into the calculation by dispersing its value.

They need to disperse it in some way that is not purely arbitrary – a way that makes some sense. Let us start with the simplest way. Let us treat a person's expectation of life as a component of her wellbeing at every time in her life; we add a component for each person's longevity into the person's wellbeing at each time. So now a person's wellbeing is more than is shown in figure 4. As well as the wellbeing shown in the diagram, each person has an extra component of wellbeing at every time in her life, given by the length of her life. This component remains constant throughout her life. What happens to our valuation of the distributions when we add in this extra component?

We first have to make snapshot valuations at each time; we have to compare the two distributions at each time separately. The two distributions in figure 3 were identical at each time, so the comparison between them was easy. But the two distributions in figure 4 are not identical at each time, so there is a further complication. In each distribution, two people are alive at each time, but they do not have the same wellbeing once we take longevity into account. In order to compare the distributions at each time, we have to aggregate together the wellbeings of the two people who are alive at that time. I have not yet said how we might do that.

The natural way is simply to add their wellbeings together. If equality is valuable, adding may not be the right way to aggregate across people. But I have set aside the value of equality, and I know no other objection to adding. So that is how I shall proceed.

We have two components of the people's wellbeings to add together, at each time. First we add up their wellbeings apart from longevity, which are shown in figure 4. Here, we find both distributions have the same total, so there is nothing to choose between them on that account. Then we add the second component, longevity. It turns out there is also nothing to choose between the distributions on grounds of total longevity either. The two people alive at each time in 4A each have a life expectancy of two years. So their total life expectancy is four years. The two people alive in 4B have life expectancies of, respectively, one and three years. Their total life expectancy is four years too. At each time, then, total life expectancy is the same in both distributions.

The conclusion is that the two distributions are equally good at every time, even taking longevity into account. Consequently, when we come to aggregate the snapshot values across times to reach an overall evaluation, the two distributions will be equally good overall. We reach this conclusion even when we try to disperse the value of longevity across times.

We shall also reach the same conclusion if we disperse the value of longevity in a different way. I have been taking the length of a person's life as a component of her wellbeing throughout her life. This component remains constant throughout her life. An alternative is to take her *future* life expectancy as a component of her wellbeing. This too is a way of dispersing value that makes some sense. The result is more complicated, because the aggregate of people's future life expectancies varies over time. (The calculation is easy to do, and I shall not spell it out.) Still, the same conclusion emerges when we aggregate across time. This sort of dispersion also gives no reason to favour 4A over 4B.

Yet if longevity is valuable, 4A is better than 4B. So dispersion seems unable to capture adequately the value of longevity. If we try to distribute the value of longevity to times, so as to make it consistent with separability, we will get the wrong conclusions about value.

This is not just bad luck. We can see the source of the difficulty from the example. Longer-lived people live for longer. Therefore, if we make snapshot valuation at individual times, each longer-lived person appears in our valuations more often than a shorter-lived person does. So longer-lived people get counted more often in the snapshot route to aggregation. This gives the result an incorrect bias. I do not see how the strategy of dispersion can get around this problem.

The bias only arises when some people live longer lives than others. This explains why inequality of lifetimes cannot be eliminated from my example.

So I think dispersion will be unsuccessful in rescuing separability of times. The only way we could justify separability is by embracing the implausible doctrine of separatism. This means denying the value of longevity. Rather than that, we shall have to abandon the snapshot approach to aggregation.

7. Causal dating

If longevity is valuable, times are not separable. Consequently, trying to value a country's health at any particular time, taking account of longevity, is doomed to failure. The reason, to put it in a nutshell, is that the benefit of living longer cannot be pinned down to a date. Indeed, the conclusion I have reached goes further than this. We simply cannot assume times are separable at all, in any context where the lengths of people's lives are not fixed. Separability of times is a bad assumption, in measuring health or in any other context.

This is a pity. We might very naturally want to trace a country's progress in combatting disease or improving life expectancy. To do this, it would be nice to have a sequence of dated valuations, to see how they change over time. It turns out that snapshot valuations are not available to make up such a sequence. However, there may be an alternative way of making dated valuations; we might attach a date to harms and benefits in a different way. Because times are not separable, we cannot always identify a date when a harm or benefit *occurs*. However, we may be able to date it according to when it is *caused*.

The cause of a harm or benefit may occur at a different date from the harm or benefit it causes. A leak of radiation may do no harm to anyone for many years, but in due course people may suffer from it. The cause is the leak, which occurs at some date; the harm occurs at a later date. It may also happen that a cause can be given a date, whereas the harm or benefit it causes cannot be tied down to any particular date. That will be so if a cause extends or shortens people's lives. We know by now that the good or harm of extending or shortening lives cannot itself be dated. But the cause of this good or harm might be datable. For example, the cause might be an epidemic or a vaccination programme, which occurs at an identifiable date.

This gives us the idea I call 'causal dating': we date harms and benefits by the date of their causes. For example, we might use causal dating for the 'burden' that disease imposes on the population of a country – the harm it does. We might calculate a sequence of dated burdens. Each would be a valuation of the harm caused by disease-in-a-particular-year. It would not be the-harm-in-a-particular-year caused by disease.

Let us think of disease as the cause and ill-health as the effect. So causal dating seems natural if we are measuring the burden of disease, whereas it would be less natural in a measure of health. Christopher Murray seems to have causal dating in mind when he proposes 'disability-adjusted life years' (dalys) as a measure of the burden of disease.[8] Take an example. Suppose an infectious disease strikes a person and leaves her disabled, The harm done by the disease continues through all the years the person's disability lasts, but the cause of the harm should be dated to the time when the disease strikes. This is how Murray dates it,[9] and this is why I think he is aiming at causal dating. I think this is an appropriate aim.

However, he does not put this aim into effect consistently, as another example shows. Suppose an infectious disease leaves a person disabled, and also weakens her constitution so that she eventually dies prematurely. Both the harm of the disability and the harm of the premature death should be dated to the time of the disease. But Murray would date the harm of the death to the date when the death itself occurs.

Indeed, I doubt that causal dating even *could* be consistently put into effect. Causal dating is not the subject of this paper, and I cannot investigate it thoroughly here. But I shall mention serious difficulties that I see with it. Nothing I have said so far stands against it, because it does not require separability of times, but it has different difficulties of its own. They are not in the realm of ethics, like those that defeat snapshot valuations. They are metaphysical problems associated with causation. It is hard to connect causes and effects in the way causal dating requires, when several causes operate. I shall explain why.

At first sight, you might think the benefit or harm caused by an event is the difference between the value of what happens given that the event occurs and the value of what would have happened had the event not occurred. Let us call this difference the ‘counterfactual’ measure of the harm or benefit. Suppose we want to evaluate the harm caused by an epidemic. We would apply the counterfactual measure as follows. Given that the epidemic occurs, there will be a particular distribution of wellbeing in the country. If the epidemic had not occurred, there would have been a different distribution. This gives us a comparison to make like the one shown in figure 2: a comparison between two two-dimensional distributions of wellbeing. We know we cannot make this comparison through the snapshot approach; we cannot compare the distributions time by time. Nevertheless, each distribution has some overall value. In principle at least, we should be able to compare their values by some other method than the snapshot one. The difference between their values is the counterfactual measure of the harm caused by the epidemic.

That is how we might try to identify the harm caused by an epidemic. But actually this counterfactual measure may not be correct. Implicitly, it takes for granted a counterfactual analysis of causation, which has well-known problems. One is the problem of overdetermination. Sometimes an effect that is actually caused by one event would have been caused by a different event had the first event not occurred. Then the effect is said to be overdetermined. The counterfactual measure of harms will not correctly identify overdetermined harms.

Here is an example. Suppose an epidemic strikes a country in two successive years. An epidemic of this particular sort always kills one tenth of the population. Suppose the country’s population is initially 1,000,000. Suppose no births occur in these years, and no deaths from causes other than the epidemics. The first epidemic kills 100,000 people, leaving 900,000 alive when the second epidemic strikes. The second epidemic then kills 90,000 people. The total number of deaths is 190,000. If the first epidemic had not occurred, 1,000,000 people would have survived into the second year. 100,000 would then have died. So the difference between the actual number of deaths and the number there would have been had the first epidemic not occurred is 90,000. This is the counterfactual measure of the harm caused by the first epidemic. However, the number of deaths caused by the first epidemic is 100,000, not 90,000. The counterfactual measure is incorrect.

As it happens, the counterfactual measure of the deaths caused by the second epidemic is correct: 90,000. Adding together the counterfactual measures for the two epidemics gives us 180,000 deaths altogether, whereas there were actually 190,000. So the problem with the counterfactual measure is not merely that it allocates effects to the wrong causes. It also miscalculates the total of effects. If we tried to use it in this example for the causal dating of harms, we would not simply give the wrong date to particular harms. We would get the wrong total of harms.

In this particular example, our intuitions about causation can correct the error of the counterfactual measure; we know which deaths were caused by the two epidemics. So we might succeed in causal dating of the harms by other means. But in other examples intuition fails too. For example, it fails when there is joint causation. Suppose an epidemic in one year kills nobody, but weakens the population. Suppose that, consequently, a second epidemic in the next year kills 100,000, whereas without the first epidemic it would have killed no one. The 100,000 deaths are caused by the two epidemics together, but we have no intuitive guidance in allocating them to the two separately. Consequently, we have no guidance in dividing the deaths between the two years, for the purposes of causal dating.

I do not know how causal dating could be applied to this second example. The

counterfactual measure is useless. If the first epidemic had not occurred, there would have been no deaths. Actually there are 100,000. So according to the counterfactual measure, the first epidemic kills 100,000 people. So does the second, for parallel reasons. Both these conclusions are evidently false.

In practice, diseases often act as joint causes, and their effects are often overdetermined. Just one example: a persisting diseases may cause gradually increasing disability in a person and eventually kill her after many years. I do not know how all the harms caused by this disease – the death and the various levels of disability – should be assigned a date of causation. Difficulties like this make me doubt that causal dating is a viable method in practice for monitoring changes in the burden of disease.

8. Decision making

I have concluded that, once the lengths of people's lives are in question, we cannot construct a dated sequence of valuations for the state of a country's health. How much of a loss is this?

It is not so bad. We can still value and compare distributions of wellbeing. We cannot do it by the snapshot approach; we cannot evaluate the state of a country's health in snapshot fashion, at a particular time. But we can still evaluate two-dimensional distributions – distributions of wellbeing across people and across time – as a whole.

I said in section 2 that separability of lives is plausible. This gives us the people approach as a feasible route to aggregation: first aggregate across times within each person's life; then aggregate across people. I do not suggest aggregation is easy by the people approach. But at least it is not ruled out from the start, as the snapshot approach is ruled out. There may be other feasible approaches too.

If it works, an alternative route to aggregation – the people approach or some other – will allow us to value and compare distributions as a whole. This is all we need for the practical purpose of making decisions. Suppose we have to decide some practical question, perhaps whether to institute a new health programme or what level of carbon tax to impose. In principle, we can evaluate the two-dimensional distributions of wellbeing that will result from each of the available alternatives, and determine which is the best. That is a sufficient basis for the decision. If we institute some programme, we cannot break down the harms and benefits it does into a sequence of harms and benefits at particular times. But to decide the value of the programme, we do not need this breakdown. Dated valuations are not needed in practice.

For instance, take the first of my epidemic examples from section 7. Suppose we are wondering whether to spend some money in order to prevent the first of the two epidemics. What is the benefit of doing so? It is the difference between the number of deaths that will take place if the epidemic occurs and the number that will take place if the epidemic does not occur. This number is 90,000. It is what I called in section 7 the 'counterfactual measure' of the epidemic's effects. I explained in section 7 that the counterfactual measure is actually incorrect as a measure of the effects, because the epidemic actually causes 100,000 deaths. Nevertheless, the counterfactual measure is exactly what is needed for decision making: the benefit of preventing the first epidemic is indeed 90,000 lives saved.

The difficulties raised in this paper are difficulties for the dating of harms and benefits. They mean we cannot present the progress of a country's health in a particular manner that would be attractive and easily to understand. But they need not interfere with our practical decision making.

Appendix: A note on separability

Separability is a formal notion that has various applications. Economists often use it in a way that is different from mine. This appendix aims to avert misunderstanding amongst economists, by adding some more explanation and some notation. In particular, it distinguishes my assumption that times are separable from a quite different assumption that may be given the same name.

If a person p is alive at a time t, let g_p^t be her wellbeing at that time, which I take to be represented by a real number. If p is not alive at t, let g_p^t have a non-numerical value that I shall designate as Ω. Ω is simply a notational device that indicates a person's non-existence at a particular time. Since g_p^t does not always stand for wellbeing, I call it p's 'condition' at t.

In section 1, I assumed that the overall value of a distribution depends only on people's conditions at all times. That is to say:

$$G = v(g_1^1, g_1^2, \ldots g_1^T, g_2^1, g_2^2, \ldots g_2^T, \ldots g_P^1, g_P^2, \ldots g_P^T),$$

where G is the overall value of the distribution, which is represented by a real number. (In this section, assume for simplicity that we are dealing with a finite number of people and times.)

My assumption that times are separable is the assumption that this function can be expressed in the form:

$$G = \bar{v}(v^1(g_1^1, g_2^1, \ldots g_P^1), v^2(g_1^2, g_2^2, \ldots g_P^2), \ldots v^T(g_1^T, g_2^T, \ldots g_P^T)),$$

where each function $v^t(\bullet)$ is real valued and $\bar{v}(\bullet)$ is increasing in all its arguments. To make the nature of the assumption more explicit, we might call it 'intertemporal separability of wellbeing' or more accurately 'intertemporal separability of conditions'.

The assumption that lives are separable is the assumption that the function can be expressed in the form:

$$G = \underline{v}(v_1(g_1^1, g_1^2, \ldots g_1^T), v_2(g_2^1, g_2^2, \ldots g_2^T), \ldots v_P(g_P^1, g_P^2, \ldots g_P^T)),$$

where each function $v_p(\bullet)$ is real valued except possibly if p never exists, and $\underline{v}(\bullet)$ is increasing in all its arguments.[10]

Those are formalizations of the assumptions I mentioned in section 2. Now I need to distinguish a different assumption. It plays no part in this paper, but it must be distinguished from those that do.

A person's wellbeing at each time she is alive will be determined by events of various sorts: the diseases that afflict her, the medical care she receives, her education, the health of her family, the food she eats, and so on. Economists often think of the relevant events as 'consumptions'. In truth, this category is too narrow, since people's wellbeing is determined by much more than just consumptions. However, because it is convenient for what I need to say, I shall treat all relevant events as consumptions too. So I assume for the sake of argument that a person's wellbeing at each time she is alive will be determined by consumptions. But for the moment I do not assume it is determined only by her own consumptions at that particular time.

Unlike an event in general, a consumption has the convenient feature that it belongs to a person at a time. Let c_p^t be the vector of consumptions of person p at time t. (Presumably it will be a zero vector at times when the person does not exist.) I am assuming that each

person's wellbeing at each time she is alive is determined by consumptions. I have also assumed that overall value G is determined by each person's condition at each time. So G is determined by each person's consumption at each time. That is:

$$G = w(c_1^{\,1}, c_1^{\,2}, \ldots c_1^{\,T}, c_2^{\,1}, c_2^{\,2}, \ldots c_2^{\,T}, \ldots c_P^{\,1}, c_P^{\,2}, \ldots c_P^{\,T}).$$

An assumption that might be called 'intertemporal separability of consumptions' is:

$$G = w(w^1(c_1^{\,1}, c_2^{\,1}, \ldots c_P^{\,1}), w^2(c_1^{\,2}, c_2^{\,2}, \ldots c_P^{\,2}), \ldots w^T(c_1^{\,T}, c_2^{\,T}, \ldots c_P^{\,T})),$$

This is obviously false. In determining people's wellbeing, there are obviously strong interactions between consumptions at different times. The value of your reading a book at some time is influenced by your education; the value to you of a tetanus injection depends on when you previously had one; and so on. The assumption of intertemporal separability of consumptions is too implausible to be taken seriously. That is why it plays no part in this paper.

But intertemporal separability of conditions is not so obviously false; it deserves to be taken seriously. This paper aims to show it is indeed false, nevertheless. In the text, I call it simply 'separability of times'. In this appendix I have distinguished it from intertemporal separability of consumption.

References

Broome, John, *Weighing Goods*, Blackwell, 1991.

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

Murray, C. J. L., 'Quantifying the burden of disease: the technical basis for disability-adjusted life years', *Bulletin of the World Health Organization*, 72 (1994), pp. 429–45.

Parfit, Derek, *Reasons and Persons*, Oxford University Press, 1984.

Notes

1. The problem of aggregating good is discussed in my *Weighing Goods* and in my forthcoming *Weighing Lives*.
2. Technically, it is *weak* separability. Separability and its implications are explained in *Weighing Goods*, chapter 4. See also the appendix to this paper.
3. I make a distinction between separability of lives and a stronger condition I call separability of people. The distinction is explained in *Weighing Lives*, chapter 6, but is not important for this paper.
4. *Weighing Lives*, chapter 6.
5. To be sure, there are objections to it, which are discussed in *Weighing Goods*, chapters 8 and 9.
6. *Weighing Goods*, chapter 11, contains a discussion of a related strong assumption of temporal separability.
7. See the discussion of dispersion in *Weighing Goods*, pp. 191–2.
8. See his 'Quantifying the burden of disease'.
9. See the section on 'Incidence versus prevalence perspectives' on pp. 431–2.
10. Separability of people is the same assumption without the qualification that v_p need not be real when p never exists. See note 3.