Journal of Public Economics 9 (1978) 91–100. © North-Holland Publishing Company

TRYING TO VALUE A LIFE

John BROOME
Birkbeck College, London W1P 1PA, England

Received January 1975, revised version received July 1977

1.

Many of the enterprises governments engage in cause people's deaths in one way or another. Therefore, those governments which like to give some of their actions the appearance of economic ratonality have to fix on a monetary value for a human life. A blossoming literature explains to them the correct way to make this valuation. But, though it blossoms still, I think that the roots of the literature are insecure. Indeed, one of my objects in this paper is to show that the attempt to value life in terms of money is more or less doomed to failure.

Before the discussion can get under way, the whole notion of the 'value' of benefits and costs, values which can be added and subtracted to establish the rightness or wrongness of a government's project, needs to be supplied both with an exact definition and a justification. Various methods of doing this have been taken up by different protagonists. For the sake of definiteness we must fix on one of them for the time being, and for that purpose I have selected the familiar interpretation of cost–benefit analysis as a 'compensation test'. Some alternatives are considered in section 2 of this paper. The compensation test interpretation defines the value of costs and benefits by means of monetary compensation, in a manner I shall specify in a moment, and then offers a justification of the definition by arguing that, if the benefits of a project exceed its costs when valued accordingly, the project is a good thing. I must explain.

I think that by now most people recognise the invalidity of the compensation test except when compensation is actually paid, so let us assume that it is. Then the test boils down to something rather harmless, the 'Pareto criterion,' and it works as follows. Each person who would lose from the project is asked what is the minimum payment of money he would consider full compensation for his loss. This amount is defined as the value of his loss. If the project is carried out it will be paid him, so he is deemed not to have suffered in the end. Each person who would benefit is asked how much money he would be prepared to pay to get the benefit. This amount is defined as its value. Then, if the value

of the benefits adds up to more than the value of the losses, the project can be carried out, the gainers can compensate the losers, and no one will end up worse off than when he started. Thus, provided some surplus is left over for somebody, the project is a good thing by the Pareto criterion.

So, from this standpoint the value of a loss is the amount of money that would compensate the loser. Coming back to the subject of death, let us for simplicity confine out attention to cases where the death in question is to be immediate, and where no bequests are permitted. Then the monetary value of a person's life, to be destroyed by a putative project, must be infinite. For no finite amount of money could compensate a person for the loss of his life, simply because money is no good to him when he is dead. There is nothing esoteric about this; it is an application, if an unexpected one, of the very orthodox notion of value which I have outlined.

But if a death counts as an infinite cost, measured in money, then it seems that a cost-benefit analysis will automatically reject any project which causes anybody's death (except possibly one which also saves lives). That, however, cannot be right. One can imagine a project which is very beneficial to millions of people, saving great suffering, but which is bound to kill somebody during the course of its implementation. Such a project would not automatically be wrong. It cannot be right for a single death to outweigh every other consideration

So there is a paradox: the theoretically prescribed way to value death appears to lead to an obviously wrong conlusion. I believe that the first person to acknowledge the difficulty in print was E.J. Mishan (1969, chs. 22 and 23), and in trying to resolve it he employed a particular device, which has been taken for granted in the flourishing literature ever since. Since I think the idea derives originally from Mishan, I shall concentrate my discussion on his arguments. Indeed, it is because he interprets cost–benefit analysis by means of the compensation test that I have done the same. Mishan's intention is to circumvent the paradox by reducing the value of death to some finite amount. The device he and his successors employ is as follows. Notice, first of all, that on the whole government projects do not kill definite people who are identifiable at the time of launching the project. Instead they increase the risk of dying for a number of people. The monetary value of the increased risk is infinite, in the sense that people will accept a finite amount of money as full compensation for it. The compensation which would satisfy the people exposed to extra risk is added up, and this is what is counted as the cost of the project. So, as it were, for purposes of evaluation death is commuted to risk of death.

The main aim of my paper is simply to show that this device is illegitimate.

To begin with, it is obvious at once that it cannot be a complete solution to the difficulty. For one can imagine an extremely beneficial project which would involve the certain death of a known person. Even according to Mishan the monetary cost of this would be infinite, but it would nevertheless be wrong

to reject such a project automatically. So the paradox is still there in that case, unresolved. I shall show that Mishan's idea is mistaken also in cases when the victim is *not* definitely known.

If a definite number of people are going to die, can it really make such a vast difference whether or not it is known who they are? Here are some intermediate cases.

(a) It is not known today who will die as a result of the project, but it is known that the information will be available tomorrow. The project is up for consideration today. It is known today that tomorrow it would be deemed unacceptable by an infinite margin, since the people who would die would accept no finite compensation. Yet today everyone has only a probability of dying. They may be compensated for this by a finite amount that turns out to be less than the benefits. So the project is accepted.

(b) The names of those who will die have been deposited by a computer in a sealed box. Had the box been opened, the project would have been rejected at once. But it has not, so the project is approved.

(c) The names are known to the government, but the people have not been told, so they can still be compensated cheaply for the extra risk. (To put flesh on these 3 examples, suppose the proposal is to close down the only treatment centre for a rare disease. The disease is fatal if not treated. Diagnosis is only possible by tests in a government laboratory, and some results are awaited.)

(d) For another example, imagine two alternative projects which have to be compared. One will cause the death of a single person, and we know who it will be. The other will cause the death of a thousand people, but it is not known which thousand. According to the proposed method, the second project is better. Yet this kills a thousand people instead of one.

(e) There may be a very small degree of probability which people do not distinguish from no chance at all, so they will accept that chance of being killed without requiring any compensation. Suppose the probability is one in ten million. Then a country with a population of fifty million could kill five people at random, assessing their deaths at no cost at all.

These examples help us to see how Mishan has gone wrong. In (d), for instance, the fact is that a thousand particular people are going to die. It is mere *ignorance* that they do not know who they are. In (b) the ignorance is plainly removable; it is being unfairly used by the government. The unfairness is especially clear in (c). Generally with future events we think that some, if not all, of our ignorance about them could be removed by looking more closely at the present. If it was wrong to use ignorance as it was used in (b), it is also wrong to use ignorance which could be removed in other ways, as by further research. But, going further, it can hardly be relevant whether the ignorance is at present removable or not. In (a) we need only wait a day. In every case it is undeniable that particular people who are alive now will die, and there is no adequate compensation for that. The government, if it follows Mishan's suggestion, really seems to be

playing a trick on people's ignorance. Provided it can get in and make the decisions soon enough, before there is much information about who will die, it can get away with causing many more deaths than if it waited. Each project which causes deaths and which is nevertheless accepted is accepted in the knowledge that, were it re-evaluated later, it would be rejected as infinitely wrong. In view of these examples, it does not seem correct to distinguish in value between the death of a known person and of an unknown person.

To put the argument more directly: if the justification for accepting a project by cost–benefit analysis is that compensation can be arranged so that nobody is harmed, then the justification cannot possibly apply when, after the project has been carried out and the utmost has been done by way of compensation, somebody palpably has been harmed, namely the person who has died.

Mishan says (p. 172):

> A word on the deficiencies of information available to each person concerning the degree of risk involved. These deficiencies necessarily contribute to the discrepancies experienced by people between anticipations and realized satisfactions. For all that, in determining whether a potential Pareto improvement has been met, economists are generally agreed – as a canon of faith, as a political tenet, or as an act of expediency – to accept the dictum that each person knows his own interest best. If, therefore, the economist is told that person A is indifferent between not assuming a particular risk and assuming it along with a sum of money, V, then on the Pareto principle the sum V has to be accepted as the relevant cost of his being exposed to that risk. It may well be the case that, owing either to deficient information, or congenital optimism, person A consistently overestimates his chance of survival. But once the dictum is accepted, as indeed it is in economists' appraisals of allocative efficiency, cost–benefit analysis has to accept V as the only relevant magnitude – this being the sum chosen by A in awareness of his relative ignorance.

This passage is remarkable in many ways. I want only to note, though, that the issue is not whether people know accurately the probability of dying, but whether they know if they are going to die. To know a probability is only a certain sort of ignorance. If people know only the probability of their dying, then the compensation they demand is chosen out of ignorance. We are asked by Mishan to accept the dictum that each person knows his own interest best. We may know our own interest better than other people, but since we do not know the future, we necessarily do not know our own interest accurately. There are some people who will die as a result of the project. Their interest is to refuse every offer of compensation, but they do not know this.

It is often said in defence of proffered techniques for evaluating projects that, although deficient, no better method is available. Something like that seems to be argued in this passage by Mishan: people's decisions are not

perfect because of ignorance, but nothing better can be done. No doubt the argument has its applications, but here it is factually in error. It is not true that nothing better can be done. The government can estimate the monetary value of the deaths much more accurately and much more easily than it can be estimated by asking people what compensation they require for the extra risk. Consider any project in which an unknown person will die. Because whoever it is does not know it will be him, because of his ignorance, he is prepared to accept a ridiculously low compensation for letting the project go forward. The government does not know who will be killed either, but it knows it will be someone, and it knows that, whoever it is, no finite amount of compensation would be adequate for him. The cost of the project must therefore be infinite, and it is only the ignorance of the person destined to die that prevents his demanding an infinite compensation. It may be true that sometimes we are forced to make decisions based on imperfect knowledge if nothing better can be done. But this is one case where the problems of imperfect knowledge can easily be eliminated. If there is to be a death, we know at once that the cost, defined as the compensation required for the loss, is infinite. Any other conclusion is a deliberate and unfair use of people's ignorance.

A way of summarising what I have argued is this. A valuation of a project may be made before it is carried out and before the distribution of its costs and benefits is exactly known, on the basis of people's choices about the risks involved. Call this an '*ex ante*' valuation. An '*ex post*' valuation, on the other hand is one made at the time of the implementation of the project, when the details of all its effects are settled. The two will often be different. My claim is that, of the two, the *ex post* valuation is the correct one (in so far as any cost–benefit analysis is correct) because it is the valuation of the actual project, whereas the other is really a valuation of the expectations created by the project. The *ex ante* valuation is useful only to the extent that it approximates to the *ex post* valuation. But, in the particular case of a project causing deaths, it is no sort of approximation at all, since the former has finite costs and the latter infinite ones. The *ex ante* valuation, in the case of death, is worthless, and furthermore it can be known to be worthless at the time it is made.

2.

That concludes my main argument. However, I started with the paradox that if life has an infinite monetary value then the false conclusion seems to be entailed that any project which causes death has to be rejected. I have now blocked the most familiar path around the paradox, so it seems appropriate to say something about how it may be resolved in other ways. As a matter of fact there is really no serious difficulty to resolve. What I have to say will seem commonplace and obvious to most people, though some may like to be reminded

of it. I shall have to deal separately with two alternative theories of cost–benefit analysis.

First of all, let us see what can be done for people who interpret cost–benefit analysis in the way I have treated it so far, as a compensation test. If, for some project, the value of the benefits exceeds the value of the costs, then compensation can be fixed up so that nobody loses. Thus, if the compensation is paid, the project is a good thing. If, on the other hand, the value of the costs exceeds the value of the benefits, then there is no way of arranging compensation so that nobody is harmed. This, unsurprisingly, is the case for all projects which kill somebody, be he known or unknown. But there is not the least reason to suppose that such projects are necessarily wrong. A project which damages some people's interests while promoting others' could well be, on balance, an improvement. The Pareto criterion simply does not make a judgment in such a case. Some people seem to have believed that, if a project is good, then necessarily it would be possible to arrange compensation so as to make nobody a loser. But there is, so far as I can tell, no warrant for this belief.

There is, then, really no paradox at all. For a project which causes a death, the costs will exceed the benefits, if the calculations are done in money. But that is no reason to reject it. A compensation test can establish that a project should be done (provided the compensation is paid), but not that it should not be done. This means that such a test can never be used to evaluate any project which causes death.

For an analogy, imagine trying to perform a compensation test with roses as medium instead of money. People cannot be compensated with roses for any major loss. Therefore, according to this method, rather a lot of projects would have an infinite cost. Nevertheless many of them could still be improvements (as we might be able to find out by recalculating their values in terms of money). The point is that roses are an inadequate measure for big costs and benefits. Money is a more powerful measuring instrument, but even the measuring rod of money is not long enough to encompass life and death. I hope this analogy will serve as a reminder that I have made no fancy claim that the value of life is infinite, but simply pointed out a difficulty in measuring it in monetary terms. Let us suppose that no finite number of roses could compensate a person for enduring a day of rain; no one would deduce that a day's fine weather is infinitely valuable.

There is an alternative interpretation of cost–benefit analysis according to which it is a test of whether the 'social welfare function' is increased or decreased. As I shall explain shortly, adherents of this view may also be inclined to value a loss by the compensation needed to offset it, so everything I have said in section 1 is equally applicable to them. We must see how these people, too, may be helped to avoid the paradox which looms as a consequence. I have to discuss this position, but in fairness to myself I hope the reader will understand that it is very far from being my own.

There is one preliminary to be clear about. People sometimes treat the social welfare function as representing what the government in fact does, or the outcome of the political system, or something of that sort. It is, naturally, possible for the government to value death in any way it wishes, and for voting or some other political process to come up with any sort of result. We are not talking about that, but rather the *correct* way to value death. So for us the social welfare function must stand for what should, correctly, be maximised. People who believe that a social welfare function can play such a role are utilitarians or members of some derivative faction. I see no point in indulging in a general argument about these functions here, nor in discussing any except an ordinary utilitarian one. (Anyone who wants to may easily apply what I say to a more general form of the social welfare function, provided it is 'individualistic.') So let

$$W = \sum_j U_j,$$

where W is 'social welfare' and U_j is the 'utility' of the jth person. U_j is a function of the state of the world. It is assumed both that the jth person in making choices chooses the alternative that maximises U_j, and also that the value of U_j represents his welfare. It is assumed that the right thing for the government to do is to maximise W, the sum of the welfare of each person.

I shall treat U_j as a function of X, the state of the world in all aspects except people's holdings of money, and of m_j, the amount of money owned by j. So $U_j = U_j(X,m_j)$. (Implicitly I assume that U_j is unaffected by anyone else's money.) Let the state of the world now be X^0, with money holdings $m_1^0, m_2^0, \ldots$. Suppose the putative project will bring about state X^1 with money holdings $m_1^1, m_2^1, \ldots$. Then the idea is that the project should be accepted if

$$\sum_j U_j(X^0, m_j^0) < \sum_j U_j(X^1, m_j^1)$$

and rejected otherwise. Let us take that for granted.

The point of the theory of project evaluation which I am describing is to find a money measure of the change in j's welfare, $U_j(X^1,m_j^1) - U_j(X^0,m_j^0)$. If amounts of money could be found for each person exactly proportional to the difference in utility, then the project would be acceptable if and only if these amounts summed up to something positive. One candidate for such a monetary measure is the amount of money which would exactly compensate j for the change (strictly, minus this amount). That is, the amount m_j^c such that

$$U_j(X^1, m_j^1 - m_j^c) = U_j(X^0, m_j^0). \tag{1}$$

We know, from the definition of integration, that

$$U_j(X^1, m_j^1) - U_j(X^1, m_j^1 - m_j^c) = \int_{m_j^1 - m_j^c}^{m_j^1} \frac{\partial}{\partial m_j} U_j(X^1, m_j) \mathrm{d}m_j$$

$$= m_j^c \frac{1}{m_j^c} \int_{m_j^1 - m_j^c}^{m_j^1} \frac{\partial}{\partial m_j} U_j(X^1, m_j)\, \mathrm{d}m_j. \tag{2}$$

The expression

$$\frac{1}{m_j^c} \int_{m_j^1 - m_j^c}^{m_j^1} \frac{\partial}{\partial m_j} U_j(X^1, m_j)\, \mathrm{d}m_j$$

is the marginal utility of money for j, averaged over values of his money holdings between $(m_j^1 - m_j^c)$ and m_j^1, when the state of the world is X^1. Call it, for short, V_j^1. Then from (2),

$$U_j(X^1, m_j^1) - U_j(X^0, m_j^0) = U_j(X^1, m_j^1 - m_j^c) + m_j^c V_j^1 - U_j(X^0, m_j^0)$$

$$= m_j^c V_j^1,$$

which follows from (1). Therefore, if m_j^c is to measure j's increase in utility, it must first be multiplied ('weighted') by an appropriate average of j's marginal utility of money. Thus, to value costs and benefits directly by the money which will compensate for them is to make, according to this utilitarian view, the tacit assumption that everybody's marginal utility of money is the same.

Before going any further, I have to mention a difficulty that afflicts me in trying to explain the application of the social welfare function in the context of this paper. For reasons separate from my own argument, the use of the social welfare function comes to seem extraordinarily odd when it has to deal with death. It becomes necessary to speak of the utility of someone who is dead, and to count it into social welfare. One of the standard objections to utilitarianism is exactly that it has this queer implication, as though there is a state of non-existence where people live, and have utility, when they happen not to exist. Richard Brautigan has the same delusion:

When you take your pill
it's like a mine disaster.
I think of all the people
lost inside of you.

I have already disclaimed any personal association with social welfare functions, and I should like to reiterate that. But in order to allow me to explain the position at all I must ask readers to swallow this curious notion. It may ease the discomfort to imagine the evaluation taking place, not when the victim is already lying dead, but just as the chopper falls.

Now, let the project under consideration kill the jth person. So in state of the world X^1 person j is dead or about to die. For this he cannot be compensated so m_j^c is infinitely negative. He cannot be compensated because money is worthless to him, but this simply means that his marginal utility of money, V_j^1, is zero. So, if it were properly weighted by marginal utility the cost of his death might be some finite sum. Then the paradox with which this paper started would vanish. Seen this way, it is nothing more than a dramatic illustration of the need to apply appropriate weights to monetary costs and benefits.

Now, multiplying the infinite compensation by the zero marginal utility is not a convenient way of arriving at a determinate value for the loss of a life. But the utilitarians have another string to their bow. So far, a person's gain or loss has been valued by the compensation which would, coming after the change, restore him to his initial well-being (called 'the compensating variation'). Instead, it may be valued by the amount of money which, coming before the change, would have the same effect on his well-being as the change itself ('the equivalent variation'), that is, m_j^d such that

$$U_j(X^0, m_j^0 + m_j^d) = U^1(X^1, m_j^1).$$

For a person whom the project proposes to kill, m_j^d is (minus) the amount of money which, taken away from him, will leave him with just the same welfare as if he were dead. The idea is conceptually staggering, but some people might claim to make sense of it, and they might suppose m_j^d to be finite. To become a proper measure of the change in utility, m_j^d needs weighting by an average marginal utility, precisely by

$$\frac{1}{m_j^d} \int_{m_j^0}^{m_j^0 + m_j^d} \frac{\partial}{\partial m_j} U_j(X^0, m_j)\, \mathrm{d}m_j.$$

This is presumably positive, being an average marginal utility for a person who is alive.

There is here a theoretically possible approach to the evaluation of death. It is beset by the general problems of utilitarianism, by the less general problems of its quantitative application in economics and by the special problem I mentioned that afflicts utilitarianism when it tries to deal with life and death. Also, it holds out little more promise of coming up with a definite figure for

the value of death than does multiplying zero by infinity. But it is at least the correct approach within its own framework.

3.

In this paper I claim to have demonstrated two things. Firstly, if an attempt is to be made to fix a monetary value on life, it is quite wrong to do it on the basis of people's evaluations of *probabilities* of death. Secondly, because the monetary compensation required for loss of life is infinite, cost–benefit analysis will be inapplicable for judging any proposal involving deaths. There is one exception to this last point. A finite monetary valuation for life could in theory be obtained by taking the equivalent variation, as opposed to the compensating variation, and weighting it by a suitable marginal utility. This option, however is open only to someone who believes that the general good can really be represented by a social welfare function, and who also attaches some sense to the utility of being dead. Moreover, there seems to be no real possibility of putting the idea into practice.

Reference

Mishan, E.J., 1969, Cost–benefit analysis (North-Holland, Amsterdam).