

PUNISHMENT.

WHAT IS 'FAIR COMPENSATION' FOR DEATH OR INJURY?

DAVID FRIEDMAN

University of California, Los Angeles CA 90024, USA

I. INTRODUCTION

The obvious principle for determining fair compensation in injury suits is that the injurer must 'make good' the damage to the injuree; that is to say, that he must make him as well off as if the injury had not occurred.¹ This principle seems to fit both our intuitions about justice and the economic view of the legal process, according to which such suits provide potential injurers with appropriate incentives not to impose risks on others, by making the injurers bear the cost of their acts.²

Unfortunately, there are some injuries for which the application of that principle seems difficult or impossible. There may be no payment large enough to make a blinded man as well off as if the injury had not occurred. If the injury is fatal the problem is inherently insoluble, save for those altruistic victims who would have been willing to trade their lives for a sufficiently large post mortem donation to their favorite charity. Even where it is possible, by enormous payments, to just barely recompense the victim of personal injury, it is not immediately obvious that the creation of blind billionaires living in profligate luxury (and so compensating, with pleasures that large amounts of money will buy, for the lost pleasures that it cannot buy) is a good idea.

The purpose of this article is to argue that 'full compensation' in the sense just described would in fact be overcompensation, in terms both of justice and of economic efficiency, and that there is another criterion which leads to more appropriate levels of compensation and is better able to deal with the 'impossible' problem of fairly compensating the victim of death or serious injury. Part II is a verbal sketch of the arguments and the criterion they lead to; in Part III I introduce and employ the formal concepts of Von Neumann-Morgenstern utility and economic rationality to redo the argument in a more precise way. Part IV summarizes the results and discusses the relation of my arguments and conclusions to conventional legal principles.

I have tried to put my arguments in a form accessible to both economists and non-economists interested in the law; I apologize in advance to those economists who find my explanations of elementary economic concepts superfluous and to those non-economists who find them insufficient.

II. THE ARGUMENT

Compensation for death or bodily injury involves two quite different problems. The

first is the problem of 'how much' damage there is to make up for. The second is the problem of in what coin damages can be paid. One might imagine that someone would be willing to give his life in exchange for a sufficiently high price—five years of ecstasy, perhaps. Faust, after all, traded not merely life but eternal bliss for a finite payment. More mundanely, we observe that people are willing to enter dangerous professions (driving dynamite trucks, for example) in exchange for somewhat higher pay, thus in effect trading life—a small increase in the probability of getting killed—for income. Both examples suggest that the reason it is impossible to 'fully compensate' someone for the loss of his life is not that the value of his life to him is infinite—it is not—but that the value of compensation to a corpse is in most cases small.³ The same argument applies to less-than-lethal injuries. If an injury somehow made the victim blind half the time—on alternate weeks—it might be possible to compensate him by a payment of a million dollars. It does not follow that two million, or even three million, would be fair compensation for making someone blind all of the time. In part this is because being blind all of the time is more than twice as bad as being blind half the time. But in part it is because in the first case, some of the payment would be spent in forms of consumption (travel, opera) which require sight to be of much value; with the possibility of such forms of consumption eliminated, larger sums must be spent on other and less attractive forms of consumption in order to provide the same value to the victim. Another way of putting the same point is to say that bodily injury makes the victim worse off in two ways. It lowers his effective income by reducing his earning power and imposing costs (a seeing eye dog, hospital bills, etc.). In addition, it lowers the value to him of any given income by eliminating ways in which he can spend it. Death is the extreme case; not only does it lower the victim's income to zero, it simultaneously reduces to zero the benefit he can get by spending any form of income—including damage payments.

One thing this argument suggests is that 'full compensation'—a level of payment for damages which restores the victim to the level of welfare he had before the injury—is in a sense inefficient. To see this, imagine that the victim first receives compensation sufficient to make up for any loss of income from his injury, so that he can afford to consume exactly the same things as if he had not been injured. Since he is no longer able to consume some of those things (color television if he has been blinded), he spends what previously went for goods he can no longer use on the remaining sorts of consumption. Since he is spending more than before his accident on these (say, gourmet dinners) one would expect the value to him of additional expenditures on these goods to be lower than before. Before his injury, the last dollar spent on 'color television services' provided the same benefit as the last dollar spent on gourmet dinners; had that not been the case, he could (and would) have improved his welfare by spending more on the one form of consumption and less on the other. After his injury, he must transfer the money he previously would have spent buying a television to buying more (and increasingly less pleasurable) gourmet dinners; it is for this reason—because he is less able to make use of income—that he is worse off even if his income is not reduced. But this also implies that with the same income, the benefit he receives from the 'marginal' dollar is less. If his compensation is sufficient to both replace his lost income and provide enough additional income to compensate for the loss of his vision (supposing that to be possible), the conclusion holds a fortiori. Since in order to make up for all the pleasures he can no longer enjoy he must consume pleasures he can enjoy to a point of near satiation, the value to him of each additional dollar will be very low indeed.

Hence 'full compensation' involves transferring income from uninjured persons, who can receive large benefits from each dollar, to injured (and already partly compensated) persons who receive very small benefits from each dollar.⁴ Intuitively it seems that although this may be fair it is also wasteful; it is doubtful that actual courts and juries (at least in cases where the injuror is a visible human being and not a corporation or insurance company) make any attempt to push compensation that far.

The reader may suspect that I am now going to argue that the proper level of compensation is that which restores the injured party to his previous level of income without making up for the losses due to his lessened ability to make use of that income. If so he is wrong. In terms of the 'efficiency' argument I have just sketched even that level seems too high, since even at that level the value of a dollar to the injured party is, I have argued, less than to a similar uninjured person with the same income. On the other hand, that level of compensation is inadequate as a deterrent to the injuror, since it understates the cost imposed by his actions.

In order to resolve this puzzle and find a level of compensation which is both efficient and adequate, we must move from the ex post situation, in which the injury has already occurred, to the ex ante situation, in which the potential victims are subject to some probability of injury or death, but the damage has not yet occurred. Two things are worth noting from this perspective. The first is that if the risk is small, there is probably some sum of money such that the potential victims would be indifferent between receiving that sum and being subject to the risk, and having neither the money nor the risk. Hence although the damage may be enormous or even infinite (in the case of death) when viewed in terms of ex post compensation, it is finite and may even be small in terms of ex ante compensation. The second and closely connected observation is that if people are going to be compensated for a deadly risk, they would much prefer to receive their money when the risk does *not* eventuate and they are therefore alive to spend it.

One possible conclusion is that those who impose risks should be required to compensate ex ante those on whom the risks are imposed. The amount of compensation would be such as to just compensate for the risk (estimated, perhaps, from the risk premium on hazardous jobs), and the potential victims could then choose how to allocate the money among the different possible outcomes. If they believed that the money was more valuable to them if the accident did not occur they could consume the ex ante damage payments; if they believed it was more valuable after the accident, they could use them to buy insurance. If they believed that some but not all of it would be needed after an accident, they could divide their expenditures between consumption and insurance accordingly.

Such a solution is unworkable. It would require the courts to estimate in advance the risks imposed by an enormous variety of activities, including some for which the very nature of the risk could not be known until too late and many for which the estimation of probabilities and potential damages would be difficult for the parties concerned and virtually impossible for the courts. What we need instead is a system under which damages are paid when an accident occurs (at which point the damage done and the fact of the accident are known) but collected (at least in part) when it does *not* occur (that being when the money is of most use to the recipient). Such a system is by no means impossible.

Imagine that we have a system in which potential victims of accidents know that they (or their heirs) will be compensated in case of injury or death; further suppose for the moment that the formula is 'full compensation'. The potential victim knows

that he is no worse off for being subject to risk; if injured he will be fully compensated. It might, however, occur to him that full compensation would involve payments of large amounts of money, much of which (in his hypothetical injured state) he could make little use of. If he were sufficiently ingenious and the society sufficiently well organized, he might then decide to *sell* insurance on himself. In exchange for some payment if he is not injured, he would agree to pay someone else part of the damage payment he receives if he is injured. Suppose, for example, that the chance of injury were one in a hundred and the 'full compensation' in case of injury were ten million dollars. He could agree, in exchange for ten thousand dollars now, to give the buyer a million dollars if he himself were injured. He would be left with nine million dollars in case of injury; while that would not be enough to fully 'make up for' the injury (if it happened he would wish it had not) the benefit he would expect to get if he 'won' his bet (i.e. was not injured) would more than make up for the loss if he 'lost' it, since a certainty of ten thousand dollars when he was uninjured was worth more to him than one chance in a hundred of having a million dollars when he was injured (and already compensated by a nine million dollar payment).

This argument implies that if potential victims are able to sell fair insurance on themselves then 'full compensation' is actually overcompensation. Prior to such a sale, the potential victim is no worse off through being exposed to risk, since, by the definition of full compensation, any damage will be made good. After the sale, the potential victim is better off than before the sale (he has transferred income from an outcome where it had low marginal utility to one where it has high marginal utility) hence also better off than if he were not exposed to risk. But if the potential injurer is overcompensating, it follows that he will be overdeterred from imposing risk; he may fail to undertake some risky activities even though the net benefit more than makes up for the risk. It further follows from this argument that the correct level of compensation is that level such that the potential victim, after selling as much insurance on himself as he wishes, will be neither better nor worse off than if no risk had been imposed. This is the same result that would follow from the system of court imposed *ex ante* damages discussed above, provided the courts had the necessary knowledge. The difference is that the injurer pays off when and only when the injury actually occurs, thus allowing the risk to be measured directly by *ex post* outcomes in the real world instead of being estimated by the court. The potential victim may then transfer income from the outcome where he is an actual victim to the outcome where he is not by selling insurance, instead of transferring it the other way by buying insurance.

While the information problems under this system are reduced, they are not eliminated. In order for the potential injurer to decide what risky actions to undertake, he must first estimate the probability and seriousness of injuries in order to calculate the damages he may have to pay. He is, unlike the court, in the best possible position to make such estimates; he is the one taking the actions and presumably the one who knows most about their consequences. In addition his welfare depends on making correct estimates; that of the court may not.

In addition to the estimate made by the potential injurer, a second estimate must be made by the insurance company which buys insurance on potential victims. Here again, the company has a private interest in doing a good job; if it overestimates the risk it will find itself paying more than the insurance is worth and losing money; if it underestimates it will be outbid by other companies with better estimates. Unlike the potential injurer, the insurance company may have no expertise in the particular

area, although it is, unlike a court, expert in the general subject of risk. And even if the estimates of the insurance companies are wrong, the result will be only a redistribution between them and their customers; the actual payments made by the injurer, and hence his incentive to avoid risky activities, will be determined by what happens, not by the insurance companies' estimates.

There remains the question of how much compensation is implied by this rule. A precise answer depends, as I will show in the next section, on the size of the risk, something which I have already argued that the court which must decide on the compensation cannot know. But as long as the risk is not very large there is an approximate answer (as I shall also show) which is independent of the size of the risk. To calculate it one simply takes the sum for which the potential victim would be willing to accept a very small probability of death (or injury) and divides by that probability. Hence if the victim were willing to accept a one in a thousand chance of death in exchange for a thousand dollars, the damages for actually killing him should be a million dollars.

Estimating the sum for which the average person would be willing to accept some small probability of death is difficult but not impossible. One approach is to estimate the wage premium on dangerous professions.⁵ The problem with this is that those who enter such professions are presumably people with abnormally low objections to lethal risks; the figure calculated in that way will accordingly underestimate the figure for the average victim. Other and more indirect ways might involve the observation of choices concerning amount and quality of medical care (where the purchaser has some grounds for estimating the effect on life expectancy), speed of driving, or other choice variables which affect probabilities of death or injury.

III. THE MATHEMATICS

A utility function is a formal way in which economists describe how the attractiveness of alternative outcomes affects decisions. It may be thought of as a numerical measure of how much an individual values various alternatives, and expressed as a numerical function of variables such as health, goods consumed, etc. The function is so constructed that if the individual prefers one outcome (consuming 50 pounds of steak a year, working 40 hours a week, and living to 90) to another (40 pounds, 35 hours, 95), then the utility of the first set of values (for the variables 'consumption of steak', 'hours worked', 'lifetime') is higher than that of the second.

It is often convenient to think of a utility function as separated into several different parts, each depending on a different variable. Thus one can think of my total utility as being made up of the utility I get from reading books plus the utility I get from playing with my children, plus . . . Such a description is a simplification of real utility functions, since preferences with regard to one form of consumption are likely to depend on how much I am consuming of something else; nonetheless the simplification is often a useful one. In considering the particular issue of injury, it may also be useful to separate effects on income from effects on consumption by imagining that the income you earn is itself a function of several inputs, among them the possession of certain innate abilities such as the ability to see. Your utility is then a function of your abilities and of other inputs, many of which can be purchased with income.⁶ The individual seeks to maximize his utility subject to a budget constraint: expenditure on consumption goods equals income (I neglect, for

a certain payment of c . Hence our ex post rule is to compensate the injured party with a payment of c/p , c having the same value as in Eq. (3).

Before going on to discuss how c/p might be estimated it is worth seeing how this solution resolves the problems raised in Section I. First, it adequately compensates the victim and deters the imposer of the risk; ex ante the victim is no worse off if the risk is imposed than if it is not. If the injury actually occurs the victim may and probably will be worse off, but if so he will have transferred some of his damage payment to the outcome in which the injury did not occur (by selling insurance on himself) and will be better off if he is not injured by an amount which, given the probabilities, just compensates him for his loss if he is.

This solution solves the problem of giving large sums of money to individuals who cannot benefit from them. Assuming that the potential victim has correctly allocated the damage payment between alternative outcomes, the marginal utility of income will be the same to him whether he is or is not injured. It also deals with most, although not all, of the situations where 'full compensation' is impossible due to the inability of the injured (or dead) victim to get sufficient utility to compensate him from any payment however large. Under this system the damages may be consumed by potential victims when they do not become actual victims and are hence able to enjoy the consumption.

There is still a potential problem. Suppose that the utility of an uninjured person is bounded; however great his income, he cannot receive a utility from it of more than $U_{\max} = U(\infty, a)$. Further suppose that the utility to him of being dead is $U(Y', a') = 0$. Lastly let the utility of being uninjured be $U(Y, a)$. If $(1-p) < U(Y, a)/U_{\max}$ it is easily seen that no possible compensation will prevent the potential victim's expected utility from being lowered by the risk. If empirical observations of the additional income which people require in order to accept hazardous jobs are to be trusted, this is likely to occur only if p is large.¹¹

To estimate c/p , the fair compensation for injury, we find some situation in which we can observe the payment (c) in exchange for which the individual is willing to accept a known probability (p') of the same injury. As I will show below, $c/p \cong c'/p'$, provided that p and p' are both small.

To see why this is the case, we first note that as long as p is small, c_3 in Eq. (3) must also be small; small amounts of money spent or received for insurance on a very unlikely event correspond to large amounts of insurance, and are therefore sufficient to cause large changes in c_4 and correspondingly large adjustments in the marginal utility of income (given the injury). The equality of marginal utility of income across outcomes then gives us:

$$U_1(Y' + c_4, a') = U_1(Y + c_3, a) \cong U_1(Y, a)$$

From this it follows that for small values of p , c_4 is approximately independent of p . (U_1 is the derivative of the utility function with regard to income.)

I now use Eq. (3) and the first order Taylor expansion of U about (Y, a) to get:

$$U(Y, a) \cong (1-p) [U(Y, a) + c_3 U_1(Y, a)] + p U(Y' + c_4, a')$$

Solving for c_3 yields:

$$c_3 \cong [p/(1-p)] [U(Y, a) - U(Y' + c_4, a')]/U_1(Y, a)$$

Substituting this into Eq. (4) gives:

$$c = (1-p)c_3 + pc_4 \cong p[U(Y, a) - U(Y' + c_4, a')]/U_1(Y, a) + pc_4$$

From which it follows that:

$$c/p \cong [U(Y, a) - U(Y' + c_4, a')]/U_1(Y, a) + c_4$$

This is approximately independent of p for p small.¹² Hence the value of c'/p' which is deduced from observed behaviour (the wage premium on risky jobs, for example) may be used as an approximate value for c/p , the amount which should be awarded to someone who suffers the same injury.

IV. THE RESULTS, THE LAW, AND THE REAL WORLD

My solution to the problem of fair compensation depends on the existence of an insurance market which allows individuals to shift compensation between different outcomes. No such market appears to exist; we see individuals purchasing insurance on themselves but not, as the argument suggests that they ought often to wish to do, selling it. One explanation is that this is a case of market failure, the transaction costs for such insurance being too great to allow the market to exist.¹³ If so, that fact increases the damages that should be paid, substituting Eq. (2) for Eq. (3).

A second alternative is that the damages currently awarded are too small, with the result that nobody wants to sell insurance on himself. Certainly the formulae sometimes used by courts to set damages in terms of an estimate of the lost income from the injury could be expected to lead to undesirably low figures.

A third alternative is that there may be legal barriers which make it difficult or impossible for an individual to sell, in advance, the damages which he will receive if injured or killed. Since I am an economist and not a lawyer, that is a possibility about which my readers may know more than I do. Certainly if such barriers do exist, my analysis suggests that they should be eliminated.¹⁴

Even if my hypothetical insurance market existed, it might be argued that it would be unfair to limit damage payments to that sum which would be fair on the assumption that the victim had made use of that market to buy or sell insurance on himself in order to minimize the ex ante cost of the risk. I would reply that this corresponds to the familiar legal doctrine according to which the liability of the injurer is limited to the damage that would have occurred if the victim had taken reasonable steps to minimize it; if the injury caused by your negligence is multiplied by my refusal to accept treatment, you are not responsible for the result. I admit that my application of that doctrine is a somewhat unconventional one.

I will end this paper by saying what I believe that I have and have not done. I have provided some foundation for the idea (which is not original with me¹⁵) that damage payments for injury or death ought to be based on the payment which the individual would require to voluntarily accept a small probability of the same injury, with damage being set equal to that payment divided by that small probability. I have also shown that that formula is not precisely correct; even under the special circumstances I assumed for my analysis it somewhat understates the proper damages where the probability of the injury which has occurred is larger than the probability of the injury used to estimate the amount of damages; in the absence of those special

circumstances it understates the proper damages substantially. I have therefore not clearly established what compensation ought to be in the real world of incomplete markets, but by showing what it should be and how it might be estimated in a world of more complete markets I have, I believe, somewhat clarified the conceptual issues involved in determining fair levels of compensation.

REFERENCES AND NOTES

1. My interest in these issues was in part provoked by Posner's brief discussion of them, especially R. A. Posner, *The Economic Analysis of Law*, Little Brown (1977), pp. 149-152, which raises but does not resolve a number of the central issues of this paper.
2. Posner (*supra*, note 1, pp. 119-161, especially 143). For an extensive and in part critical discussion of the validity of the economic approach to law, see (1980) 9 J. Legal Stud.
3. This point is made explicitly in J. Broome, 'Trying to Value a Life', (1978) 9 J. Pub. Econ. 91-100 in his discussion of the utilitarian approach to the valuing of life. My approach is similar to that discussed (but not accepted) in that part of his paper, save for the crucial difference that I do and he does not introduce Von Neumann-Morgenstern utility. See also J. Broome, 'Trying to Value a Life: A Reply', (1979) 12 J. Pub. Econ. 259-262; J. M. Buchanan and R. L. Faith, 'Trying Again to Value a Life', (1979) 12 J. Pub. Econ. 245-248; M. W. Jones-Lee, 'Trying to Value a Life', (1979) 12 J. Pub. Econ. 249-256; and A. Williams, 'A Note on Trying to Value a Life', (1979) 12 J. Pub. Econ. 257-258.
4. Readers familiar with such arguments will realize that the result is not a rigorous one; it is possible to construct (bizarre) utility functions for which it does not hold.
5. See for instance Linnerooth and Bloomquist (G. Bloomquist, 'The Value of Human Life: An Empirical Perspective', (1981) 19 Econ. Inq. 1) and their references. For an interesting discussion of value of life see G. Tullock, *The Logic of the Law*, Basic Books (1971), pp. 155-158.
6. One way to increase the realism of the model, at considerable cost in simplicity, is to allow the individual to allocate a fixed budget of time (24 hours a day) between time used to earn income and time used, with consumption goods, to produce utility. This would be particularly appropriate in those cases in which the same injury that reduced income also increased leisure.
7. Strictly speaking, I will be talking about utility functions and expenditures in one or another of two different probability states ('states of the world'): one in which the injury occurs and one in which it does not. In the state in which the injury occurs, some expenditures occur before the injury. Since the individual only knows which state he is in when the injury occurs (or after the time when it could occur has passed), it is convenient, although not strictly accurate, to refer to expenditures or utility functions before or after the accident, and I shall sometimes do so.
8. Broome, *supra*, note 3, tries to discuss a utilitarian approach without introducing choice under uncertainty. He thus (deliberately or not) makes it impossible to discuss the utility of life save in his preferred context of a money payment sufficient to compensate for a certainty of immediate death. But this is impractical because of the declining marginal utility of money, a point he has already made.
9. J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behaviour*, Princeton University Press (1953), pp. 15-29.
10. Expected utility is defined as the probability of each outcome in a lottery times the utility of that outcome, summed over all outcomes. It corresponds, in terms of utility, to the expected return of a lottery in terms of money.
11. In this case, at least, Broome, *supra*, note 3 is correct in arguing that money is an insufficient yardstick for measuring life, at least as long as we restrict ourselves to costs and benefits to a single individual and do not permit interpersonal utility comparisons.

12. Readers with a taste for rigour may fill in for themselves a more precise statement of the theorem whose proof I have sketched here.
13. One source of market failure might be 'reverse moral hazard'. The individual who has sold his right to collect damages will increase the precautions he takes to prevent or minimize injury; while such action may be desirable from an efficiency standpoint it also means that what he sells has a lower expected value to the buyer than it would have had to the seller, which may prevent the sale.
14. One obstacle might be the lack of any insurable interest on the part of the buyer (I owe this point to an anonymous reviewer). While this legal doctrine may be justified in other contexts on the grounds that my ownership of an insurance policy on someone whose life is of no value to me gives me an incentive to increase its value by increasing his probability of death, it is worth noting that in this case the (desirable) incentive effects on the seller may be more important (see note 13).
15. Readers who are familiar with the literature on liability will realize that I have deliberately ignored a large number of issues (such as standards of negligence) in order to concentrate on one. Those familiar with the conceptual literature on valuing a life (in particular Broome (*supra*, note 3) and the subsequent debate) will appreciate my reasons for trying to avoid philosophical arguments about the meaning and relevance of probabilistic considerations and expected utilities wherever possible. Any serious attempt to set forth my views on the relevance of Paretian and utilitarian comparisons to issues of choice and welfare would have made this a much longer and very different paper.