HEIDELBERG UNIVERSITY
DEPARTMENT OF ECONOMICS

Conducting Cost Benefit Analysis in Expected Utility Units Using Revealed Social Preferences

David Canning

AWI DISCUSSION PAPER SERIES NO. 722
January 2023

Conducting Cost Benefit Analysis in Expected Utility Units Using Revealed Social Preferences

David Canning[1]

December 2022

Abstract

Assuming individual preferences satisfy the Von Neumann–Morgenstern axioms for expected utility we show how we can measure an individual's expected utility of any state using their willingness to accept a gamble over two reference points. The utility function captures the diminishing marginal utility of money with income and risk aversion over gambles. This contrasts with the standard money metric valuations that assume linearity of an individual's welfare in money. Measuring costs and benefits in expected utility units seems more appropriate than money units for applied welfare economics since it reflects individuals' preferences more accurately, and can be applied to policies that involve risk. In addition, if social preferences satisfy the Von Neumann–Morgenstern axioms and the Pareto principle, social welfare is the weighted sum of these expected utilities. The weights can be calculated directly for the United States from revealed Government preferences on the allocation of mortality risk. The United States Government values lives equally in calculating the welfare losses from mortality risk and this implies an equal weighting of individual utilities if they are measured using willingness to accept a gamble of a probability of death versus the status quo; we call this life metric expected utility. For projects with small effects on expected utility, we show how to convert existing money metric cost benefit studies into life metric expected utility cost benefit analysis using weights based on how the money value of a statistical life varies with income in the United States. Our approach may be particularly appealing for the conduct of cost-benefit studies mandated by regulation in the United States to inform Government policy. It measures costs and benefits in expected utility units that respect individuals' preferences over risk and sums these utility gains using the Government's revealed preferences and implied social welfare function.

Key words: Welfare economics, cost benefit analysis, value of life.
JEL Numbers: D60, D61, D63.

---

[1] Harvard T.H. Chan School of Public Health

Acknowledgement: I would like the thank the participants in the Brocher Summer Academy in Global Population Health: Healthy, Wealthy, and Wise - The Ethics of Health Valuation, Geneva, 2022, and the Third Society for Benefit-Cost Analysis European Conference, Paris, 2022, for stimulating discussions on this paper.

The author has no competing interests.

1. Introduction

Many theories of social welfare involve evaluating states based on individual utilities. With utilitarianism we simply sum the individual utilities, more complex approaches involve weighted sums of utilities, or sums of non-linear functions of the utilities (Adler 2019). The practical application of these approaches requires knowledge of the individuals' utility functions. While these utility functions are not directly observable, we can observe people's preferences through their choices and use these to infer their utility function. Assuming individual preferences satisfy the Von Neumann–Morgenstern (VNM) axioms for expected utility, we show that by asking questions about willingness to accept a standard gamble over two reference points, we can derive a person's Von Neumann–Morgenstern utility function over all possible states (Von Neumann and Morgenstern 1947). By summing the individual utility gains or losses from a project, we have a form of economic cost benefit analysis where costs and benefits are measured in utility units.

Our proposed expected utility approach has a major advantage over the use of money willingness to pay if the projects under consideration involve uncertainty. Our derived utility functions accurately reflect each individual's preferences; in particular the curvature of the utility function over money captures the diminishing marginal utility of money with income and correctly models risk aversion over gambles. Money willingness to pay has been shown to generate a money metric utility function in the case of choices over goods with no uncertainty (Weymark 1985, Fleurbaey 2011). However, money metric utility is a non-linear transform of the underlying true utility function; it is transformed so as to be linear in money, and loses the concavity and risk aversion that may be present in the original utility function (Blackorby and Donaldson 1988). Thus money metric utility does not accurately reflect preferences over gambles unless people are risk neutral. However, there is strong evidence people are not risk neutral and indeed are often risk averse (Elminejad, Havranek et al. 2022). Money metric cost benefit analysis has frequently been criticized for assuming that the marginal utility of money is constant across people in interpersonal comparisons, which is questionable on ethical grounds (Blackorby and Donaldson 1990), we emphasize the additional criticism that it assumes the marginal utility of money is constant for each person across all states in intrapersonal comparisons. While the problem of interpersonal comparisons requires ethical judgments, preferences over gambles do reveal the individual's VNM utility function and we argue that this utility function, based in willingness to pay in terms of a standard gamble, is a better guide for policy than willingness to pay measured in money.

A particularly important case where the fact that money metric approach does not accurately measure preferences over uncertainty is in valuations of states which involve a probability of death. The money compensation required to accept a probability of death is non-linear in the probability, indeed the money compensation required to accept death with certainty may well be unbounded, but expected utility by definition is linear in probabilities. Given that the compensating money variation required to impose death on a person with certainty, or a very high probability, is likely to be unbounded above, policies involving a certainty of death for any one individual are never optimal in money metric cost benefit analysis. To avoid this problem, the theory is usually only applied to projects involving small changes in the risk of death to individuals and the money value of a statistical life (MVSL) is only measured in the limit as the change in the probability of death goes to zero, and should not be applied to non-marginal changes in the probability of death (Adler 2020). Our approach to measuring costs and benefits is to use expected utility units. Expected utility is linear in the probability of death and can be applied across the full range of mortality risk. Policies that involve non-marginal changes in the probability of death, or large scale uncertainty, such a global warming, or pandemic responses, will therefore not be accurately evaluated by MVSL (Adler 2020) but will be in our expected utility approach.

It could be argued that the money metric approach can be extended to measure the money equivalent of large risks of death. However, while the compensating variation required to accept a large risk of death is much greater than the found by applying the MVSL the willingness to pay to avoid this avoid large a risk of death is generally smaller than the MVSL due to being bounded by lifetime wealth (Hugonnier, Pelgrin et al. 2022). The large gap between willingness to pay money to avoid, and willingness to accept money to bear, a large risk of death makes money metric cost benefit analysis problematic since it is based on the assumption that the two measures are approximately equal and that the Hicks-Kaldor criteria give the same results (Hammitt 2015). For example, the Kaldor criterion classifies a policy intervention as socially beneficial if the total willingness to pay of those who gain exceeds the total willingness to accept of those who are harmed. In a status quo point where 100 million people are going to die is evaluated against an alternative policy where these people live but a single different person dies then the Kaldor criteria will reject the policy, the willingness to pay of the 100 million to live is bounded by their wealth while the compensation required to the one who dies may be unbounded. On the other hand, the Hicks criteria uses equivalent variations and the willingness to pay of the one person to avoid the policy is bounded by his wealth and will be much smaller than the amount needed to be paid to the 100 million to forgo the policy. The money metric cost benefit approach is therefore likely to be inconclusive when evaluating policies with different distributions of mortality risk. Our expected utility approach avoids this problem (see online appendix) .

The expected utility approach has been used for policy evaluation of policies involving risk, for example for deriving the optimal level of unemployment benefit financed through taxes on workers, which trades off gains from risk sharing against losses due to incentive effects (Chetty 2006). However, these studies assume that everyone has the same utility function, with same parameters, which can be calibrated based on parameter estimates from existing studies of behavior. Our approach extends this type of expected utility analysis to policy questions where individual preferences may be very diverse and need to be elicited.

Our approach to adding up the utilities we derive as a measure of social welfare is open to the critique that preferences only identify the individual's Von Neumann–Morgenstern utility function up to a positive linear transformation and we will get different results for social welfare if we multiply one person's utility by a large factor. The precise specification of the utility function we derive depends on the two reference points in the standard gamble we use to elicit preferences, the difference between the reference points defines a unit of utility. We could view the true utility function, which measures the intensity of pleasure, as well as preferences over choices, as key to implementing utilitarianism. This approach requires the direct measurement of utility or happiness (Layard, Mayraz et al. 2008). However, even if the intensities of pleasures are measurable, which is unclear (Broome 2008), this hedonistic approach falls foul of the utility monster critique (Nozick 1974); if one person is very good at turning resources into pleasure, society may want to give all its resources to that person, which seems unfair to the more stoic individuals. Instead of thinking of the utility function as measuring pleasure, we think of it merely as a representation of a person's preferences, with each representation being equally correct. In this case, the choice of normalization of each person's utility function is a normative ethical decision since it only affects interpersonal comparisons of welfare when we add utilities together and not individual preferences and choices.

If the social planner's preferences satisfy the VNM axioms and the Pareto principle then the social welfare function can be written as a weighted sum of the individual utility functions; the social welfare function is utilitarian (Harsanyi 1955, Hammond 1992). The weights can be thought of as a particular normalization of the individual VNM utility functions; arbitrary multiplication of an individual's utility function can be undone by changing the weights to leave the social welfare function unchanged. The result can therefore be restated as a simple utilitarianism, without weights, provided the reference points used to elicit utility and define utility units are chosen appropriately.

One approach would be to construct the weights, or reference points, from ethical principles but this seems fraught and is unlikely to lead to consensus. A more practical approach is to use an inverse optimal argument. If the Government is using a social welfare function in its decision making, its policy choices reveal its preferences and the utility weights it is using. These utility weights can then be used in cost benefit analysis to make its social rankings consistent with the Government's revealed preferences. It is possible to derive utility weights by revealed social preferences on the distribution of money through the tax and benefit system, including transfer payments. While this can be done (Bourguignon and Spadaro 2012, Hendren 2020), taxes have a complex structure and have incentive effects that produce an efficiency loss, and the implied social welfare weights depend on estimates of the magnitude of these distortions, which are uncertain.

We prefer to use revealed social preferences on allocating risks of death. For example, in the United States the Office of Management and Budget recommends using a common money value across people for risks of death, and the Department of Transport, and the Environmental Protection Agency, each value lives equally in deciding on projects to reduce the risk of death (Andersson and Treich 2011, Robinson 2020). The consistency of social preferences implied by the VNM axioms means that the social welfare function derived from these revealed social preferences on the distribution of mortality risk should be the same as used by the Government in all its decision making.

Valuing lives equally implies the United States Government's social welfare function normalizes all individual utilities to one at the status quo and zero at death for each individual. If lives are valued equally by society, these life metric expected utilities can simply be added together to get total social welfare without weighting. Our utility functions do not measure pleasure but rather the probability of death a person is willing to accept to achieve a state; we therefore call it life-metric utility. Our utility functions are cardinally measurable, and fully interpersonally comparable. The only indeterminacy that we have is in the normalization of the reference state to utility one and death to utility zero. However, normalization to these states to different numbers imposes the same linear transformation on all utilities, and simply shifts the utility sum by the same linear transformation, and has no effect on social rankings of states. We assume lives are valued equally over a finite current population; we do not address the issue of potential future lives and their weighting which creates difficult ethical problems (Broome 2004).

We can compare our approach to ranking projects using the sum of life metric expected utility to rankings from using money metric cost benefit analysis. In the United States it is well known that the money value of a statistical life varies with income. People in the highest income quartile are willing to pay six times as much money as a person in the lowest income quartile to avoid the same small probability of death (Kniesner, Viscusi et al. 2010). Using money metric cost benefit analysis implies that the government should be willing to spend more on reducing mortality risks for the rich than for the poor. In particular, the cost threshold for spending per statistical life saved should be higher for car safety measures than for public transport safety since car occupants tend to be richer than public transport users. However, this approach has been rejected and the United States department of transport uses the same value of a statistical life for all modes of transport. It therefore appears that in the Unites States, government agencies value lives equally at the status quo. In the United Kingdom public policy is based on valuing life years lived in full health equally across people (Chilton, Jones-Lee et al. 2020). We show how to reweight our life metric utility measures to conform to this different revealed social preference.

An objection to the practical implementation of our approach is that people are not used to thinking about willingness to pay in terms of a probability of death and may find it difficult to do so, especially for projects that involve small changes in utility. Willingness to pay in terms of probabilities of death have been used in health state valuations, but are best reserved for considering large changes in utility. We show an equivalent method of measuring utility is to measure the willingness to take a gamble over a money loss and convert this to life metric utility using the willingness to pay money to avoid a risk of death. The money gamble elicits the VNM utility function and this can transformed to life metric expected utility by rescaling.

For projects that involve only small changes in utility, we can measure the change in utility for each person by their willingness to pay for the change in money terms, weighted by their marginal life metric utility of money which is the inverse of their money value of a statistical life (MVSL) based on their willingness to pay money to avoid a small probability of death as usually measured in value of life studies (Kniesner, Viscusi et al. 2010). This linear approximation allows us to convert existing money metric cost benefit studies to life metric expected utility gains and losses. While studies of the trade-off between money and a probability of death are usually called value of life studies we prefer to think of them as value of money studies; they tell us how people value money in life metric utility units and allow us to compute the marginal utility of money.

In our life metric expected utility approach the value of a life is the same for everyone; the status quo has a utility 1 and the status quo with the person's death a utility of 0. The expected utility of death with probability p is (1-p). The money willingness to pay for a small probability of death we call the money value of a statistical life (MVSL) rather than the more common term value of a statistical life (VSL) to emphasize it is a money metric measure rather than a utility measure. In our approach the value of a statistical life in utility units is equal across all people, with a value of one, while the value of money (its marginal utility) varies across people.

We emphasize that the linear approximation to convert money to life metric utility only holds for policies that have small effects on utility. For projects that may have large effects, this approximation may be misleading, and a full elicitation of utilities using a standard gamble may be required. The major existing critique of money metric cost benefit analysis has been on ethical grounds, that it assumes that the marginal utility of money is equal across people. We add to this critique the point that it assumes the marginal utility of money is constant for individual's but this is unlikely to be true across large changes in income and utility. Money metric approaches can therefore only be used to address problems where total utility effects are small but these will, of course, be less interesting issues than policies that could have large effects on utility, such as climate change or pandemics. Once large changes are considered, we need to measure expected utility directly and we show in this paper that direct measurement is feasible.

Our work is closely related to several existing literatures. The idea that if preferences are well ordered, they can be represented by an expected utility function is well established (Von Neumann and Morgenstern 1947), and eliciting the VNM utility function is a version of the standard gamble used in health state valuation (Rowen and Brazier 2011). The health state approach uses full health and death as the reference points to evaluate other health states that lie between these extremes. Our innovation is to use the method to rank all states not just health states. We also innovate by developing a method of calculating the utility of states better than the status quo, and worse than death, that are outside the range of the reference points, so we can rank all states.

The result that if the social planner's preferences satisfy the VNM assumptions, and they obey the Pareto principle, they have social welfare function that is a weighted sum of the individual expected utilities is again well known (Harsanyi 1955, Hammond 1992). Our innovation in this setting is to derive the weights from a revealed preference argument using observed government decisions on distributing a risk of death. This is similar to the inverse optimal approach deriving social weights on money transfers from the progressivity of the tax and benefits schedule (Bourguignon and Spadaro 2012, Hendren 2020). However, the tax and benefit system is complex and the implied social weights depend on the magnitude of the incentive effects of taxation on economic efficiency which is highly uncertain. In the United States, willingness to accept a risk of death is a more attractive measuring rod for utility than willingness to pay money. The Government's revealed preference of is that social value of a risk of death is equal across people, while the social value of money is unequal. Our approach to generating the weights used in the social welfare function is to use this revealed social preference approach. We could use a time trade off approach to measure costs and benefits in healthy life year equivalents in a framework with no uncertainty (Canning 2013), but this misses the advantages of the expected utility approach that allows for choice under uncertainty.

A paradox in the MVSL literature has been that while the MVSL is known to vary widely across people, for policy purposes a common MVSL is used when evaluating mortality risks, valuing lives equally, which is inconstant with cost-benefit analysis which emphasizes the use of private valuations. A weighted money metric approach is compatible with valuing lives equally if the social weights on money are the inverse of each individual's value of statistical life, their willingness to pay to avoid a risk of death (Somanathan 2006, Baker, Chilton et al. 2008, Farrow 2021). We show that this is a first order approximation to our life metric expected utility, valid for small changes in utility. An advantage of our life metric expected utility approach is that it applies over the full range of outcomes not just marginal changes. We also emphasize a change in interpretation of the equivalence result. Rather than thinking of valuing lives equally as being the result of a special weighting of money metric utility across people (Somanathan 2006, Baker, Chilton et al. 2008) we take valuing lives equally as the revealed preference of the United States Government which implies these particular marginal utility weights on money are correct (Farrow 2021).

It is well known that the numeraire used for cost benefit analysis affects social rankings since the willingness to pay for non-traded goods varies across people (Brekke 1997). We not only change the numeraire but also shift to an expected utility measure so that when we add up over people we are utilitarian. It has been argued that weighted utilitarianism is incompatible with valuing lives equally (Adler, Hammitt et al. 2014). However, this is due to their assumption that social preferences must satisfy anonymity which we do not impose (we assume only that they satisfy the VNM axioms for expected utility and the Pareto principle).

An alterative way of constructing weights for the social welfare function that has been employed is to argue that individuals are indifferent between states that have only small differences in utility, and have a least noticeable difference in utility (Argenziano and Gilboa 2019). Assuming society is indifferent between reallocations that provide people with a unit of their least noticeable difference in utility we can use this to construct weights for the social welfare function. However, this approach is open to version of the utility monster critique, a connoisseur critique; we may want to redistribute to the rich if they are highly sensitive connoisseurs able to detect even small differences in utility. The authors make the argument that the rich may be more insensitive to different utility levels than the poor but this may not be so; high income and the experience of a variety of foods, housing, and entertainment, may lead to an increased appreciation of small differences and subtleties that are lost on the less experienced. Most importantly the least noticeable difference approach does not seem to be used in Government policy, which would be required in a revealed social preference framework.

In section 2, we show how to elicit the individual's VNM expected utility function based on their preferences over gambles. In section 3, we construct the social welfare function as a sum of weighted individual utilities, based on the axiom that social preferences satisfy the Pareto principle. In section 4, we show how to calculate the weights based on the revealed social preference. In the United States, where Government values all lives equally, this is sufficient to generate a unique social welfare function based on simply summing individuals' life metric expected utility. Section 5 shows how to carry out our approach using willingness to pay in terms of a gamble measured as the probability of a money loss, and how to convert results of existing cost benefit studies in money units to approximate life metric expected utility units. Section 6 concludes.

2. Eliciting the Von Neumann–Morgenstern Utility Function

We take a set of n possible outcomes S_j for j=1, … , z as given and define simple lotteries over these outcomes as a set of non-negative probabilities that sum to one. For example, the simple lottery A would be $p_1S_1 + p_2S_2 + ... + p_zS_z,\ p_j \geq 0, \sum_j p_j = 1$. We allow for lotteries that are probability distributions over simple lotteries and lotteries over these and so on; these more complex lotteries can be multiplied through to give probability distributions over states. For simplicity we take the set of possible outcomes to be finite; generalizing the results to lotteries that are probability measures over a Borel measurable space with a sigma-algebra is straightforward (Hammond 1992).

We assume that individuals have preferences over lotteries given by $A \succ B$ if A is preferred to B and $A \sim B$ if the individual is indifferent between A and B. We write $A \succsim B$ to mean one of $A \succ B$ or $A \sim B$ holds. We assume that the social planner also has preferences over lotteries.

We assume both individual and the social planner's preferences satisfy the following four axioms:

Axiom 1 (Completeness) For any lotteries A*, B*, exactly one of the following holds:
$A \succ B, A \sim B, or, B \succ A$.

Axiom 2 (Transitivity) If $A \succ B$ and $B \succ C$, then $A \succ C$, and similarly for $\sim$.

Axiom 3 (Continuity): If $A \succ B \succ C$, then there exists a p such that

$$pA + (1-p)C \sim B.$$

where the notation on the left side refers to a situation in which A is received with probability *p* and C is received with probability (1–*p*).

Axiom 4 (Independence): For any C and $p \in (0,1]$,

$$A \succ B \Leftrightarrow pA + (1-p)C \succ pB + (1-p)C$$

Given axioms 1 through 4, the expected utility theorem (Von Neumann and Morgenstern 1947) implies there exists a von Neumann–Morgenstern utility function u from the set of states to the set of real numbers such that

$A \succ B \Leftrightarrow E[u(A)] > E[u(\mathrm{B})]$

Where $E[u(A)] = p_1 u(\mathrm{S}_1) + p_2 u(\mathrm{S}_2) + ... + p_z u(\mathrm{S}_z)$

Any positive linear transformation of the function u is also a VNM utility function for the individual and gives the same preferences over lotteries as u.

In social choice theory it is often assumed that the social planner knows the utility functions of the individuals. More precisely in the case where agents have preferences and VNM utility functions defined up to a positive linear transformation the social planner knows the family of VNM utility functions that represent an individual's preferences. In our case these utility functions are cardinal but non-comparable across individuals (Sen 1986). Rather than assume the social planner knows of the individual utility function directly we assume that the social planner can elicit an individual's preferences between states.

Axiom 5 Non trivial domain there exist a lottery R and a lottery D with the property $R \succ D$.

Axiom 6 Preference elicitation. For any two states A and B a social planner can elicit if $A \succ B, A \sim B, or, B \succ A$ for an individual. For any three states $A \succ B \succ C$, the social planner can elicit the probability p such that $pA + (1-p)C \sim B$.

A difficulty with eliciting preferences over states is that each state is a complete description of the current and future world (indeed universe). Such complex states will be difficult to describe and may even be indescribable due to the complexity involved. A possible feasible alternative is to take the status quo point as given and understood by individuals and define reference points in terms of a finite number of differences from the status quo. The first part of axiom 6 implicitly assumes that the social planner can describe states to individuals in a way that they can understand.

The fact that there exists a probability that makes the individual indifferent between B and a lottery over A and C does not make it obvious that the social planner can elicit this. One issue is that the probability is a real number while any probability stated to a finite number of decimal digits will only be an approximation to this. In addition, if the elicited preferences are used for policy purposes, revealing their true preferences may not be incentive compatible for individuals. The second part of axiom 6 assumes these issues can be overcome. We now give an algorithm the social planner can use to construct the individual's VNM utility function.

Theorem 1: Utility Elicitation The social planner can elicit the individual's utility $u(X)$ of any lottery, where $u(X)$ is a VNM utility function representing the individual's preferences.

By the expected utility theorem axioms 1-4 guarantee the existence on a VNM utility function v for the individual. We now define a measurable function $u(X)$ over all possible lotteries X.
We begin by asking the individual to rank X relative to states R and D with $R \succ D$. By axiom 5 such states exist. By axiom 1 this ranking is possible and by axiom 2 the results are transitive.

Hence by axioms 1,2 and 5 there are 5 mutually exclusive possibilities of the ranking of X relative to the reference states.

(i) If $X \succ R$ we define p by the property $(1-p)X + pD \sim R$. By axiom 3 such a p exists and by axiom 6 it can be elicited. By the expected utility theorem we have $(1-p)v(\mathrm{X}) + pv(D) = v(R)$. Now define $u(X) = \frac{1}{1-p}$. Hence, we have $u(X) = \frac{v(X) - v(\mathrm{D})}{v(\mathrm{R}) - v(D)}$ which exists since $R \succ D \Rightarrow v(R) > v(D)$

(ii) If $X \sim R$ we have $v(X) = v(R)$ and we define $u(X) = 1 = \frac{v(X) - v(\mathrm{D})}{v(\mathrm{R}) - v(D)}$

(iii) If $D \prec X \prec R$ we define p by the property $(1-p)\mathrm{R} + pD \sim X$. Again, this p exists and can be elicited by axioms 3 and 6. And again by the expected utility theorem we have $(1-p)v(\mathrm{R}) + pv(D) = v(X)$
In this case we define $u(X) = 1-p$ and we have $u(X) = \frac{v(X) - v(\mathrm{D})}{v(\mathrm{R}) - v(D)}$

(iv) If $X \sim D$ we have $v(X) = v(\mathrm{D})$ and we define $u(X) = 0 = \frac{v(X) - v(\mathrm{D})}{v(\mathrm{R}) - v(D)}$

(v) If $X \prec D$ we define p by the property $(1-p)\mathrm{R} + pX \sim D$. Again this p exists by axiom 3, and again by the expected utility theorem we have $(1-p)v(\mathrm{R}) + pv(X) = v(D)$. We now define $u(X) = \frac{-(1-p)}{p}$ and we have $u(X) = \frac{v(X) - v(\mathrm{D})}{v(\mathrm{R}) - v(D)}$.

Hence for every X we have

$$u(X) = \frac{v(X) - v(\mathrm{D})}{v(\mathrm{R}) - v(D)} = \frac{1}{v(\mathrm{R}) - v(D)} v(X) - \frac{v(\mathrm{D})}{v(\mathrm{R}) - v(D)}$$

Therefore $u(X)$ can be constructed by questions over preferences and is a positive linear transformation of the VNM utility function $v(X)$ since $R \succ D \Rightarrow v(R) > v(D)$ and hence $u(X)$ is a VNM utility function for the individual.

QED

The actual utility function we elicit depends on the choice of reference points, different reference points generate different affine transformations of the utility function. Suppose we use the reference points $C \succ E$ to elicit an alternative utility function $\tilde{u}$, then it is easy to show that we have

$$\tilde{u}(X) = \frac{1}{u(\mathrm{C}) - u(E)} u(X) + \frac{u(D) - u(E)}{u(\mathrm{C}) - u(E)}$$

The advantage of this result is that if we do want to change the reference points we do not need to elicit the utility function again, we can simply transform it to the new function provided we have elicited the utility of the new reference points in terms of the utility function defined by the old reference points. This suggests that if there is a debate about the right reference points to use in empirical work, we should collect data about the utility of alternative reference points at the reference points actually being used so that the alternative approach can be implemented if desired.

The original proof of the VNM theorem (Von Neumann and Morgenstern 1947) only used standard gambles in the cases (ii), (iii) and (iv) to measure utility. The utility of any outcome is measured in terms of a gamble between more extreme outcomes, one on either side. Our approach to using revealed social preferences to calibrate the VNM utility function uses observed preferences over a fixed set of reference points and it may be that the utility of an outcome we are interested in lies outside this range, particularly if the range of possible utilities is unbounded. We therefore allow outcomes outside the range of utilities covered by the reference points.

Our approach to measuring utility is similar to the approach in the literature on valuing health states, where the reference points are taken as full health and death (Torrance 1986). Case (iii) in theorem 1 corresponds to valuing health states that lie between full health and death and case (v) corresponds to ranking states that are worse than death. In the health utility approach full health has a utility of 1 and death a utility of 0, and there is discomfort about the fact that the utility of a state worse than death can be unbounded below (Robinson and Spencer 2006). Negative utilities in health are therefore transformed to be bounded by -1, however we do not do this since the resulting transformed number loses its interpretation as a utility. The health literature does not consider health states better than full health and so does not deal with case (i). Our interpretation of utility is that it is a numeric representation of people's preferences over outcomes ex ante before they know the state that occurs, not a measure of pleasure associated with a state. This has the advantage that when we consider preferences on lotteries that include a risk of death we do not have to imagine the dead actually having a utility function ex post, which is somewhat problematic (Becker and Stout 1992).

The questions asked and the probabilities elicited in the different cases in theorem 1 are very different depending on if the utility state to be ranked is in the range of the reference points or outside. For example, in case (i) the probability of the lower reference point D is the compensating variation required at the new outcome X to make the individual indifferent to this and R. In case (iii) the probability of the lower reference point D is the equivalent variation where the individual is indifferent to this and getting the outcome X. The interpretation of the probability p in each case is different and the way it is transformed into a utility measure is different.

3. Social Welfare

We assume there are a finite number of individuals, n, and that the social planner has preferences over states that satisfy the axioms 1-4. We now subscript the preference ordering to denote whose it is, $\underset{i}{\succ}$ is the ordering for individual i while $\underset{W}{\succ}$ is the ordering for the social planner.

The fact that the social planner's preferences satisfy axioms 1 to 4 implies the social planner has a VNM utility function representing these preferences. We also assume the social preference ordering satisfies the Pareto Principle.

Axiom 7 Strong Pareto Principle

If for all i we have $A \underset{i}{\sim} B$ then $A \underset{W}{\sim} B$

If $A \underset{j}{\succ} B$ for some person j and $A \underset{i}{\succsim} B$ for all $i \neq j$ then $A \underset{W}{\succ} B$

Axiom 8. Revealed Social Preferences.

There exists a reference point R and n lotteries, D_i , one for each person i, such that for all j and all $i \neq j$

$$D_i \underset{W}{\sim} D_j$$

$$\mathrm{R} \underset{i}{\succ} \mathrm{D}_i$$

$$\mathrm{R} \underset{j}{\sim} \mathrm{D}_i$$

In order to identify the weights (or reference points for the standard gamble) to use when we add utilities together, we need to make the simplifying assumption that there exists a private good for each person (sometimes called independent prospects). A private good is one that affects one person's utility but does not affect the utility of anyone else. Given the existence of a private good, we can use revealed societal preferences about the distribution of this good to determine the weights being used in the social welfare function.

In order to construct the weights for the social welfare function we need the social planner to reveal their preferences on this private good. In money metric cost benefit analysis it is usually assumed that money is a private good for each person so that money payments affect the utility of the person paying but not the welfare of other people. If a person's money payments affected the welfare of others, we might have to compensate the others for the effects of these payments, and so on, ad infinitum. People's preferences, and hence utility functions, over social states include not only the effect of the state on themselves but also its effect on others; people may prefer states in which others are better off, or states in which others are worse off. To some extent, people include a social preference ordering in their personal preferences but in a utilitarian approach we want a measuring rod for each person's utility that reflects their personal preferences measured in a way devoid of their social preference.

To avoid the unreasonable assumption that there actually exist private goods, we could imagine an extended set of possible social states that allows for each person to experience a different state but leaving everyone else's experience unchanged (Milleron 1972). That is, we ask a person to rank different social states they will experience on the assumption that everyone else stays at the status quo state. In this way we can measure individual preferences about how states affect the individual, not how people rank states depending on how they also affect others. If we do assume a real private good that does not affect the utility of other people, it may be better to think of our utility functions as representing household preferences rather than individual preferences since preferences within households are highly interdependent.

We take the case in which a private bad has to be distributed across society, we could equally well take the issue of distributing a private good. Given a reference point R we imagine the social planner has to give a private bad to someone, for example a probability of death, or a loss of money. For example we could take D_i to be the reference point R but with a probability that individual i dies immediately, all else being unchanged. The first part of axiom 8 means the probabilities of the worse outcome for each person in the lotteries are such that the social planner is indifferent about which of these lotteries takes place. The second part of axiom 8 assumes that each individual i prefers the reference point to the worse outcome. The third part of the axiom is that the worse outcome for each individual i is a private good so that each person j is indifferent between the reference point and the worse outcome for person i, $i \neq j$.

An important point is that we assume the social preference over distribution of the private good is revealed only at the reference point R. We do not need to know social planner's preferences at other points.

Theorem 2 the Social Welfare Function

The social welfare of a state X is given by

$$W(X) = \sum_i^n u_i(X)$$

Where u_i is the VNM utility function of individual i elicited at the reference points (R, D_i).

Proof.

The social planner has preferences that satisfy the VNM axioms and Pareto indifference. Hence by Harsanyi's theorem (Harsanyi 1955, Hammond 1992) the social planner has a social welfare function that can be written as

$$W(X) = \sum_{i}^{n} w_i u_i(X)$$

where $A \underset{W}{\succ} B \Leftrightarrow W(A) > W(B), \text{and}, A \underset{W}{\sim} B \Leftrightarrow W(A) = W(B)$

Our assumption of the strong Pareto principle further ensures that all the weights are strictly positive since a zero or negative weight on any individual will violate the principle.

By axiom 8 we have $D_j \underset{W}{\sim} D_k$ for arbitrary j and k hence

$$W(\mathrm{D}_j) = \mathrm{W}(\mathrm{D}_k) \Rightarrow \sum_{i}^{n} w_i u_i(\mathrm{D}_j) = \sum_{i}^{n} w_i u_i(\mathrm{D}_k)$$

But by axiom 8 we also have for $i \neq k, i \neq j$ that $u_i(D_j) = u_i(D_k) = u_i(\mathrm{R})$

Hence we have

$$\sum_{i \neq j} w_i u_i(\mathrm{R}) + w_j u_j(\mathrm{D}_j) = \sum_{i \neq k} w_i u_i(\mathrm{R}) + w_k u_k(\mathrm{D}_k)$$

Cancelling individuals other than j and k on both sides gives

$$w_k u_k(\mathrm{R}) + w_j u_j(\mathrm{D}_j) = w_j u_j(\mathrm{R}) + w_k u_k(\mathrm{D}_k)$$

Which, since the weights are strictly positive, implies

$$\frac{w_k}{w_j} = \frac{u_j(\mathrm{R}) - u_j(\mathrm{D}_j)}{u_k(\mathrm{R}) - u_k(\mathrm{D}_k)} = 1$$

Since the utility functions using these reference points for each individual have a gap of one between their reference points by construction. Hence the weights on each individual are equal in the social welfare function if we use their utility functions measured as the stated reference points. Since the weights are equal, we can normalize them to one since the rankings implied social welfare function are invariant if all weights are changed by the same affine transformation. QED

Our social welfare function has several appealing properties. It satisfies the VNM axioms and can be interpreted as maximizing expected social welfare in cases of uncertainty. It satisfies three of the assumptions of Arrows impossibility theorem (Arrow 2012), non-dictatorship, unrestricted domain, and Pareto efficiency. It avoids the impossibility theorem by violating the irrelevance of independent alternatives assumption. Ranking of any two states depends on the choice of our reference states.

A desirable property of the social welfare function we lack is anonymity. The social welfare function maps individual preferences into social preferences and anonymity requires that a permutation of individual preference orderings among individuals does not change the social preference ordering. Under anonymity, the social preference ordering depends on the individual orderings, but not on which person has which ordering. If we permute preferences across people, our social welfare function changes, indeed it breaks down and is undefined. If individual i gets person's j preferences but keeps his reference point (R, D_i) where D_i differs from R only in terms of a private good for person i, he will be indifferent between the two outcomes since person j was indifferent between them. In this case, his utility measured using the reference point (R, D_i) is undefined. We leave to future research the issue of whether with a suitable choice of reference points some type of symmetry of treatment across individuals can be maintained.

Our construction of life-metric utility has focused specifying the individual utility functions in a way that is cardinal and fully interpersonally comparable that allows us to measure social welfare using a utilitarian social welfare function that sums individual utilities (Weymark 2016). Approaches to measuring social welfare that involve non-linear functions of individual utilities, such as prioritarian approaches, will have different information requirements. For example the Atkinson social welfare function (Atkinson 1970, Adler 2019) requires utilities that are non-negative and measured on a ratio scale which our approach does not satisfy, since we allow for states that have negative utility when they are worse than the lower reference point.

4. Revealed Social Preferences

The social welfare function we have constructed is depended on the set of reference states (R, D_i) used by the social planner which determine the precise representation of the utility function we pick from the class of VNM utility functions for the individual, which are only defined up to a affine transformation. As shown in section 2 changing the reference point is equivalent to putting different weights on each individual's utility function. One approach would be to determine these from ethical principles but this seems fraught. A more appealing approach is to think of the Government as the social planner and infer the reference points – or weights – from their revealed preferences. An advantage of this approach is that we only have to observe the Government's preferences over allocating one private good to determine its weights, and these can then be applied to all cost benefit calculations. This approach may be particularly appealing for cost benefit studies mandated by regulation to inform United States Government policy (Dudley 2020) since by using the Government's revealed utility weights the rankings using the cost benefit study will be consistent with the Government's preferences and may increase the quality and usefulness of such studies (Hahn and Dudley 2020).

Given a social welfare function, there will be an implied optimal tax system that will redistribute income from those with low marginal social utilities of money to those with high social marginal utilities, taking into account distortions due to the effect of the taxes on incentives (Atkinson and Stiglitz). Given a tax system, we can undertake the inverse optimal approach to determine what weights the government puts on redistribution of money to people at different levels of income. Hendren (2020) uses this approach to derive weights on money transfers by income level for the United States. However, this approach requires the use of empirical estimates of how distortionary taxes are. Different components of the tax and benefit system cause different levels of distortion, and there is a wide range of plausible estimates for these different levels of distortion. While Hendren finds the plausible range of distortions always imply a higher weight on money transfers on those with lower incomes, how much higher is in a range of 1.2 to 5 times the weight on high income earners.

We can see how this inverse optimal approach works in a simple example. A money metric expected utility approach could be justified if the Government used a poll tax, levying a fixed tax burden on each person independent of circumstances (Adler 2020). Consider a choice of possible poll taxes for, one for each person i, b_i which the government chooses subject to the constraint that the sum of taxes equals a fixed level of Government sending. Since a fixed poll tax does not generate distortions to efficiency we do not have to take account of these effects on the budget constraint. Now let b_i* be the poll taxes actually chosen. The Government has revealed that this choice of b_i* maximizes their social welfare function. For each individual we can construct a VNM utility function based on the reference state being the post-tax position and the worse reference state being the status quo point with a tax of $1. We can define social welfare over the vector of choices of taxes $(\mathrm{b}_1,...,\mathrm{b}_n)$ where this is understood as the status quo with the addition of this vector of lump sum taxes. By Harsanyi's theorem we have

$$W(\mathrm{b}_1,...,\mathrm{b}_n)=\sum_i^n w_i u_i(\mathrm{b}_i)$$

for some fixed weights.

Maximizing social welfare subject to the government budget constraint, assuming money is a private good for each individual, and the utility function is differentiable in money, gives the first order conditions for a maximum

$$w_i \left.\frac{du_i}{db_i}\right|_{b_i*} = w_j \left.\frac{du_j}{db_j}\right|_{b_j*}$$

for all individuals i and j, where the derivatives are evaluated at the optimal poll tax levels.
Now we have chosen to measure the utility function using the post-tax status quo as giving a utility of 1 and the state with $1 less as giving a utility of 0. The marginal utilities at the optimum are given by

$$\left.\frac{du_i}{db_i}\right|_{b_i*} \approx u_i(b_i*+1)-u_i(b_i*)=-1$$

The marginal utility cost is approximately 1 for an extra dollar of poll tax.

Hence the first order condition for the poll tax to be optimal is

$$w_i = w_j$$

Poll taxes reveal the social planners social welfare function to the sum of individual utilities measured so that the loss of $1 results in the loss of one unit of utility for each person; money metric utility. The intuition for this result is that with poll taxes, the social planner can redistribute $1 across any two people while keeping the budget constraint unchanged. At the optimum, any redistribution of this type must be undesirable, which means that the marginal social utility of money is equal across all people. The use of a poll tax would therefore imply the social welfare function should be the sum of money metric expected utilities. This is similar to the current money metric approach to cost benefit analysis, though we should elicit willingness to pay using probabilities of having a money loss (rather than direct money equivalents) to ensure we take account of the curvature of the utility function. While a poll tax would justify money metric expected utility, we would emphasize that, in practice, governments do not use poll taxes but rather prefer progressive taxes which redistribute income to the poor even though these have efficiency loses and each dollar gained by the poor costs more than a dollar of income to the rich.

Given the complexities of the incentive effects in actual tax systems, a simpler approach is to apply the inverse optimal approach to Government policy on allocating mortality risk. In the United States, evaluation for policy decisions on the mortality risks due to environmental exposures and transportation accidents values lives equally despite the fact that there are arguments that people with different incomes, health states, and ages should be valued differently (Andersson and Treich 2011, Robinson 2020). In the United States, therefore, the harm done by an environmental or transport effect on deaths is measured by the total sum of expected lives lost, independent of whose life it is; we simply add up additional probabilities of death induced by different policies. This implies that the Government is indifferent to shifting a fixed probability of death from one individual to another, since it leaves total expected deaths unchanged. In terms of axiom 8, this means taking the reference point as the status quo and the individual bad as death. If we measure utilities using these reference points, the social welfare function used by the United States Government is simply the sum of these utilities. For losses relative to the status quo, we ask people what probability of death would be equivalent to the loss, while for gains, we ask what probability of death would make them indifferent to the change and staying at the status quo. We then use the formula in section 2 to calculate the utility of the new state.

While in the United States the dominant approach in public policy involving mortality risk is to value all lives equally, in the United Kingdom the approach is to value healthy life years equally. This means that in the United Kingdom the death of a person lowers social welfare more if they are healthier or have a longer life expectancy if they were not to die. However, the value of a life is not adjusted for other factors such as income. This suggests that for the United Kingdom we can set the higher reference point as the status quo but with full health and a fixed life expectancy, of say 20 years, with the low reference point as being dead. If we measure utility in life metric units, which is appropriate in the United States, these utilities have to be weighted by the value of the person's health and longevity, relative to the norm of full health for 20 years, measured by their willingness to accept a probability of death to get the better outcome.

5. Cost Benefit Analysis in Units of Life Metric Expected Utility

We now turn to the issue of implementing cost benefit studies using life metric utility which is appropriate for the United States. We also address the question of converting existing cost benefit and cost effectiveness studies into life metric utility terms. It could be argued that people are not used to evaluating their willingness to pay in terms of a probability of death, but are used to thinking of willingness to pay in terms of money. It may be easier for respondent to answer questions in units they are familiar with and then transform the answers to the desired numeraire (Carthy, Chilton et al. 1998). However, in section 2 we have shown that we can easily convert VNM utility in one metric to another using a linear transformation. Suppose we measure utility in money terms based on a loss \$m to each person, that is we scale each persons utility function to have a value 1 at the status quo and a value 0 at the status quo but with a loss of \$m. However, we want to measure the utility function using a life metric expected utility where the good state is the status quo and the bad state is death. As shown in section 2, we can do this by applying the transformation

$$u^*(X) = \frac{1}{1-u^m(\text{death})} u^m(X)$$

where u* is the desired life metric utility and u^m is the money metric utility for the money loss m. We ignore the constant term in the transformation since this applies equally to all states and does not affect individual or social preferences. Death will usually be worse than the payment of a fixed lump sum \$m (our results go through in all cases but this is the most intuitive). Given the way we have constructed utility measures we have

$$u^m(\text{death}) = \frac{-(1-p(\text{m}))}{p(\text{m})}$$

where p(m) solves the equation

$$(1-p(\text{m}))(\text{status quo}) + p(\text{m})(\text{death}) \sim (\text{status quo} - \$\,\text{m})$$

Or

$$(1-p(\text{m}))u^m(\text{status quo}) + p(\text{m})u^m(\text{death}) = u^m(\text{status quo} - \$\,\text{m})$$

hence

$$u^*(X) = p(\text{m})u^m(X)$$

Life metric utility is simply a weighted money metric utility, where the weight is the individual's equivalent probability of death to the amount of money loss m used in the construction of the money metric utility. This approach is fully consistent and should give the same utility measures whether we ask for equivalent probabilities of death directly, or if we ask for equivalent probabilities of a money gamble over the amount m and then covert them using the weights p(m). Hence, as well as asking for willingness to pay in terms of a gamble involving the loss of $m versus the status quo, we propose cost benefit studies also measure the probability of death equivalent to $m so that results can be converted to life metric utility units.

If we consider only small variations around the status quo and use marginal arguments, we can go further and convert existing cost benefit studies in money units into life metric expected utility. $m is the amount of money a person would be willing to pay to avoid a risk of death of p(m). We have taken m as given and derived p(m). We could equally well take probability of death p as given and use the indifference condition to derive m(p) where $m(p)=p^{-1}(m)$.

We now define the expected utility of the status quo with a money loss $m and a probability of death p as

$$\mathrm{U(p,m)} = (1-p)u^{m}(\text{status quo}-\$\,\mathrm{m}) + pu^{m}(\text{death})$$

The equation we have used to define p(m) is exactly the same as used in the literature on estimating the money value of a statistical life (MVSL). The MVSL for an individual is defined as

$$MVSL = \lim_{p\to 0}\frac{\mathrm{m}(p)}{p} = \lim_{m\to 0}\frac{m}{p(\mathrm{m})} = \left.\frac{dU/dp}{dU/dm}\right|_{U=U_0} = \left.\frac{dm}{dp}\right|_{U=U_0}$$

The MVSL is the willingness to pay, per unit, for a small reduction in mortality risk. It can be interpreted as the relative marginal expected utilities of a probability of death and money in the expected utility function, and is the slope of the indifference curve in the probability of death and money, holding expected utility steady at U_0 the level at the status quo. Note that the ratio of marginal expected utilities is the same for any VNM utility function, since affine transformations affect both marginal utilities equally.

Hence taking $m to be one dollar and assuming this is small enough to allow the ratio to be close to the limit, we have

$$u^*(X) = p(\mathrm{m})u^{m}(X) \approx \mathrm{m}\frac{dp}{dm}u^{m}(X) = \frac{1}{MVSL}u^{m=1}(X)$$

It follows that we can derive life metric utility from money metric utility, measured so that $1 gives 1 unit of utility by multiplying the money metric utility by the inverse of the person's MVSL. People with a high MVSL reveal they have a low marginal utility of money (assuming we value lives equally) and hence the social planner should have a low weight on their money willingness to pay. It is important to note that the MVSL weights we use are the individual's MVSL.

Valuing lives equally is often thought to be at odds with the observation that the money MVSL varies widely across people, with systematic differences by income, health, and age. In a utilitarian approach, weighting money metric willingness to pay with the inverse of the MVSL, treating the marginal social utility of money as the inverse of the MVSL, is known to result in valuing lives equally (Baker, Chilton et al. 2008); the same result as we present. However, our interpretation is different. Their interpretation is that valuing lives equally would only be socially desirable if the social marginal utilities of money across people happen to be the inverse of their individual MVSLs. An better interpretation is that the United States Government has revealed its preference for valuing lives equally and therefore its social weights on money are the inverse of the individual MVSLs (Farrow 2021). Further, we can apply these weights to all public policy decisions since they completely define the Government's social welfare function. These approximations show that for projects that lead to small changes in expected utility, our approach can be considered a form of weighted money cost benefit analysis, where the weight on money willingness to pay is the individual's marginal life-metric utility of money (the inverse of their MVSL).

Estimates for MVSL by household income quantile are available for the United States (Kniesner, Viscusi et al. 2010). These are shown in Figure 1. Figure 2 shows the inverse MVSL by household income, which can be used as weights on money values in cost effectiveness studies to derive approximate life metric expected utility. A regression of the log inverse MVSL on quantile log average income gives a coefficient of -1.55, suggesting we can use the formula $MVSL^{-1}=Y^{-1.55}$ to calculate inverse MVSL weights for the United States by income level Y. The estimates in Figures 1 and 2 suggest that the Government weights a dollar for the rich (the top quantile) at about 1/6 the value of a dollar for the poor (the bottom quantile). If the Government is consistent in the welfare weights it uses for policy, then we should derive the same weighting scheme from the inverse optimal approach applied to taxes. However, the results to date imply a somewhat higher weight on the rich relative to the poor than we find, though there is a considerable range of uncertainty depending on what assumptions are made as to how distortionary are the effects of taxation (Hendren 2020).

Figure 1: Money Value of a Statistical Life (MVSL) by Income Quantile in the United States

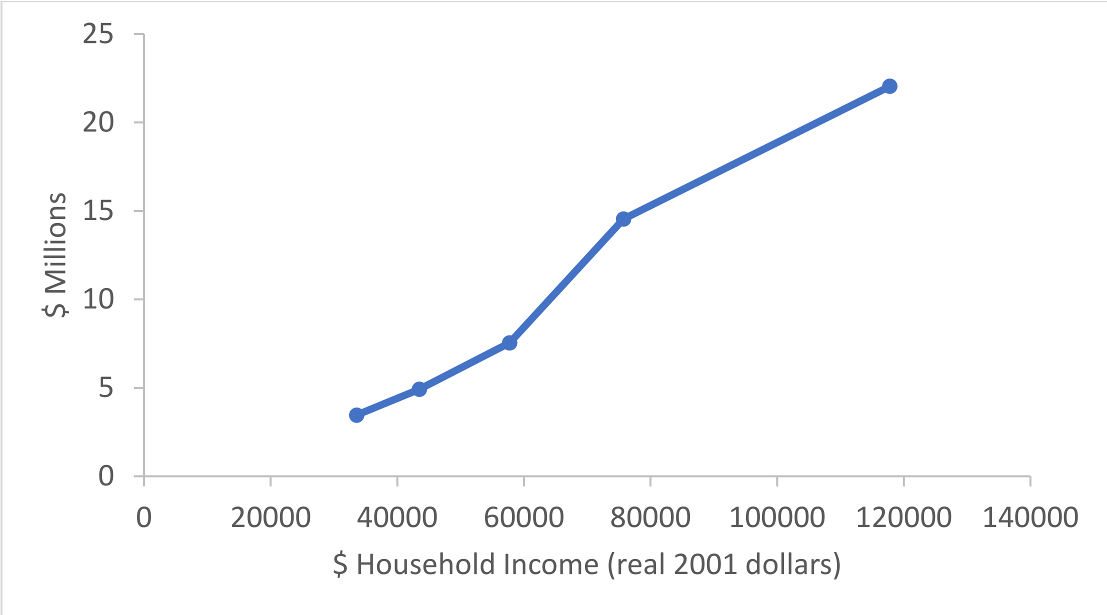
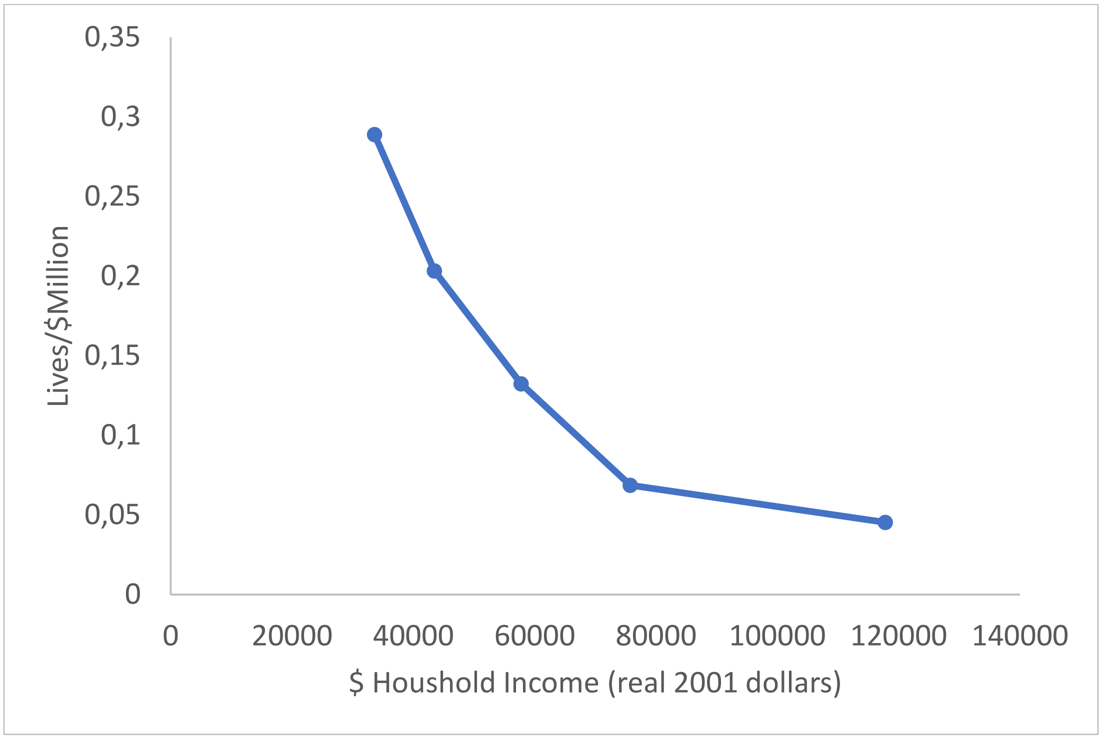


Figure 2: Inverse of the Money Value of a Statistical Life (MVSL) by Income Quantile in the United States



We would emphasize the use of variations in MVSL across people within a country. These studies tend to find an elasticity of MVSL with income greater than 1. Studies on the variation of the average MVSL across countries tend to find a much lower income elasticity, which is incompatible with the other evidence of how quickly the marginal utility on money declines with income (Kaplow 2005). However, these cross-country studies may suffer from aggregation bias since the average MVSL in a country is not the same as the MVSL of a person at its average level of income.

VSL has been shown to vary substantially with income in the United states (Evans and Schaur 2010, Kniesner, Viscusi et al. 2010) and in Taiwan (Liou 2018) . The MVSL has also been shown to vary with age and baseline health (Aldy and Viscusi 2020, Herrera-Araujo and Rochaix 2020). In existing cost benefit studies that use money units, we can treat the absence of MVSL measures as a missing data problem and impute the unobserved money value of a statistical life for a person from that of people with similar characteristics in existing MVSL studies, using multiple imputation to correct for the additional uncertainty this generates in life metric expected utility (Murray 2018). An important point is that we need to adjust both money benefits and money costs to utility units. Often the distribution of money costs is not reported in money metric cost benefit studies since results only depend on total costs. Utility losses from costs, however, do depend on the distribution of money costs. We can approximate this distribution of costs by the distribution of the tax burden across income groups for projects financed out of general taxation.

In the United Kingdom where healthy life years, rather than lives, are valued equally, the utility function has to be further adjusted. The reference points for measuring utility should be a life lived in full health for say 20 years and death, rather than the status quo health and survival profiles and death as in the United States. Again, if we do not have utilities measured using these reference points, we can convert life metric utility to quality adjusted life year metric utility by using the individual's willingness to pay, in terms of accepting a probability of death, to get a future life in full health for 20 years. These values can also be imputed based on existing studies. For example, we can use studies of health utility (Rowen and Brazier 2011) which measures the utility of health states in terms of willingness to take a standard gamble involving a risk of death, which is exactly what we require for our approach. Note that this valuation of health states should individual specific, different people may value health states differently, and we should not impose average values for the population, rather we should impute based on the person characteristics, such as income level, health state, and age. We can also convert existing cost effectiveness studies in the United Kingdom to expected utility cost benefit analysis measured in units of quality adjusted life years. The health benefits are already measured in the correct units of expected quality adjusted life years gained. However, money costs of health interventions should be weighted to reflect the marginal utility of money of those paying, which will involve dividing the money payments by the inverse of the individual's wiliness to pay money units for a life year in full heath.

Table 1 shows the relationship between our utility metric approach and other approaches to economic evaluation. Currently the dominant approach is cost benefit analysis which measures costs and benefits in money units. Dissatisfaction with this approach for health projects led to the development of cost effectiveness analysis where costs are still measured in money units, but benefits are measured in health units. Cost utility analysis has been applied in the health field to transform health benefits into utility units using a standard gamble approach. A difficulty with the cost effectiveness and cost utility approaches is that the costs and benefits are measured in different units and so the net effect of a project is unclear. Our approach is to measure both costs and benefits in utility units, for all costs and benefits, not just health outcomes, and add these up to get the effect on welfare defined as the sum of utilities.

Table 1. Economic evaluation methods and units

Type of Analysis	Cost Unit	Benefit Unit
Cost Benefit	Money	Money
Cost Effectiveness	Money	Health
Cost Utility	Money	Utility
Utility	Utility	Utility

6. Conclusion

Our proposed changes to cost benefit analysis has two major elements, the first is that we should measure costs and benefits in expected utility units, the second is that we should add up these utility costs and benefits using weights derived from revealed social preferences. The first, we do not think should be controversial. Indeed, the persistence of money metric cost benefit studies is somewhat odd given the ease with which expected utility can be measured and the superiority of the expected utility approach in economic theory, particularly for the evaluation of situations where the curvature of the utility function matters, such as in attitudes to risk. In the health literature, it is common to measure the utility of heath states using a standard gamble between full health and death, and we simply advocate a similar approach to measuring all states for economic evaluation.

When measuring the expected utility of a state, we need to choose two reference points, with different utility levels, to use a measuring rod. These two points define a "unit" of utility. Any two points will give the same Von-Neumann Morgenstern expected utility function, defined up to a linear transformation. Different reference points simply give different normalizations. The reference points may however be chosen so as to be easy for respondent to interpret in different situations; such as a gamble between the status quo and a money loss.

Our second point is that if the social planner has a social welfare function that is a weighted sum of the individual expected utilities, then social policies will maximize this function. Given this, we can use an inverse optimal approach on observed policies to infer the social weights being used. This approach has been used to infer social weights from tax schedules. A non-distortionary uniform poll tax, independent of any personal characteristics, would imply the Government values a unit of money equally for each person. However actual tax schedules are more complex and inferring the social weights is difficult given the need to adjust for the incentive effects of taxes on behavior; the progressivity of the optimal tax system depends on both the utility weights and the efficiency costs of taxes. However, it is much easier to derive the weights being used for policies about the distribution of mortality risk. In the United States, the Government weights a probability of mortality equally across individuals for transportation and environmental risks. This implies its social welfare function is utilitarian where utilities are measured using the reference points of the status quo and death. If utilities are measured in terms of a money gamble it implies weighting these utilities by the inverse of the individual's money value of a statistical life. In the absence of utility measures based on willingness to accept a money gamble, we can use existing money metric cost benefit studies, weighted by the inverse of the individual's money value of a statistical life, as an approximation to expected social welfare for small changes in utility. Using the Government's revealed preferences in weighting utilities may be particularly appealing when carrying out cost benefit studies mandated by regulation in the United States to guide policy making.

There are many possible objections to our approach. At the individual level, the existence of well-ordered preferences that obey the VNM axioms has been questioned and other approaches proposed, though a version of our utilitarian approach may be possible even without the VNM axioms (McCarthy, Mikkola et al. 2020). People may have difficulties in calculating with probabilities which we require for the elicitation of preferences. Even if the individual can calculate the probability required the issue of incentive compatibility of revealing these preferences to the social planner may pose problems. Treating the Government as a social planner is even more problematic in the United States, where voting on policies, and the separation of powers, means actual policies chosen may not be internally consistent with a single well-ordered preference relation.

An alternative to our revealed social preference approach based on Government decisions would be to use ethical principles to decide the normalization of the utility measures, the choice of reference points, to be used to construct utilitarian social welfare. A particular concern is that our revealed social preference approach gives ethical significance to the status quo in constructing the reference points. The private good allocation of each individual may differ at the current status quo implying that social preferences violate anonymity. A symmetry of treatment axiom would be appealing but might require that the private good allocation at the reference points be the same for each person. In addition, if we allow the "status quo" to change over time, we change the social welfare function and we get the well known preference reversals and inconsistencies of cost benefit analysis applied to a sequence of different status quo points. A fixed pair of reference points, symmetrical across people, independent of the status quo, would be theoretically very appealing. But implementation of this approach would require agreement on, and a complete description of, these reference points, which seems difficult. Our revealed social preference approach makes cost benefit analysis consistent with the preferences the Government has revealed by its policies; a more modest aim than defining what social preferences should be.

Despite these difficulties, our approach has a major advantage of providing a theoretically rigorous and empirically implementable alternative to simple money metric cost benefit analysis. In future cost benefit studies using contingent valuation, we suggest that, as well as eliciting willingness to pay in money units, researchers elicit willingness to pay in terms of accepting a gamble involving a money loss. They can also elicit willingness to pay in money for a small mortality risk. These two extra questions allow the construction of money metric expected utility and life metric expected utility; and the comparison of results across the three approaches. We would emphasize the results may be very different. To return to our example from the introduction suppose policy makers in the United States have to choose between a policy where 100 million people die, or a policy where 1, different, person dies. As we have shown money metric cost benefit analysis is inconclusive in this case while in life metric expected utility the gain from the policy with only 1 death is clearly superior. While heavily criticized, money metric cost benefit analysis has persisted due in the main to the lack of a viable alternative; we argue cost benefit analysis using life metric expected utility, provides such an alternative.

References

Adler, M. D. (2019). Measuring social welfare: An introduction, Oxford University Press, USA.

Adler, M. D. (2019). "Social Welfare Functions." Global Health Priority-Setting: Beyond Cost-Effectiveness: 123.

Adler, M. D. (2020). "Benefit–cost analysis and distributional weights: an overview." Review of Environmental Economics and Policy.

Adler, M. D. (2020). "What should we spend to save lives in a pandemic? A critique of the value of statistical life." A Critique of the Value of Statistical Life (June 25, 2020). Duke Law School Public Law & Legal Theory Series(2020-40).

Adler, M. D., J. K. Hammitt and N. Treich (2014). "The social value of mortality risk reduction: VSL versus the social welfare function approach." Journal of Health Economics 35: 82-93.

Aldy, J. E. and W. K. Viscusi (2020). "Age differences in the value of statistical life: revealed preference evidence." Review of Environmental Economics and Policy.

Andersson, H. and N. Treich (2011). The value of a statistical life. A handbook of transport economics, Edward Elgar Publishing.

Argenziano, R. and I. Gilboa (2019). "Perception-theoretic foundations of weighted utilitarianism." The Economic Journal 129(620): 1511-1528.

Arrow, K. J. (2012). Social choice and individual values, Yale university press.

Atkinson, A. B. (1970). "On the measurement of inequality." Journal of economic theory 2(3): 244-263.

Baker, R., S. Chilton, M. Jones-Lee and H. Metcalf (2008). "Valuing lives equally: Defensible premise or unwarranted compromise?" Journal of Risk and Uncertainty 36(2): 125-138.

Becker, W. E. and R. A. Stout (1992). "The utility of death and wrongful death compensation." Journal of Forensic Economics 5(3): 197-208.

Blackorby, C. and D. Donaldson (1988). "Money metric utility: A harmless normalization?" Journal of Economic Theory 46(1): 120-129.

Blackorby, C. and D. Donaldson (1990). "A review article: The case against the use of the sum of compensating variations in cost-benefit analysis." Canadian Journal of Economics: 471-494.

Bourguignon, F. and A. Spadaro (2012). "Tax–benefit revealed social preferences." The Journal of Economic Inequality 10(1): 75-108.

Broome, J. (2004). "Weighing lives." OUP Catalogue.

Broome, J. (2008). Can there be a preference-based utilitarianism. Justice, political liberalism and utilitarianism: Themes from Harsanyi and Rawls. J. Broome, M. Fleurbaey, M. Salles and J. A. Weymark**:** 221-238.

Carthy, T., S. Chilton, J. Covey, L. Hopkins, M. Jones-Lee, G. Loomes, N. Pidgeon and A. Spencer (1998). "On the contingent valuation of safety and the safety of contingent valuation: Part 2-The CV/SG" chained" approach." Journal of risk and uncertainty 17(3): 187-214.

Chetty, R. (2006). "A general formula for the optimal level of social insurance." Journal of Public Economics 90(10-11): 1879-1901.

Chilton, S., M. Jones-Lee, H. Metcalf, J. S. Nielsen, R. Baker, C. Donaldson, H. Mason, N. McHugh, R. McDonald and M. Spackman (2020). "A scoping study on the valuation of risks to life and health: the monetary Value of a Life year (VOLY)."

Dudley, S. E. (2020). "Regulatory oversight and benefit-cost analysis: a historical perspective." Journal of Benefit-Cost Analysis 11(1): 62-70.

Elminejad, A., T. Havranek and Z. Irsova (2022). "Relative Risk Aversion: A Meta-Analysis."

Evans, M. F. and G. Schaur (2010). "A quantile estimation approach to identify income and age variation in the value of a statistical life." Journal of Environmental Economics and Management 59(3): 260-270.

Farrow, S. (2021). "On Balance: When All Lives Matter Equally: Equity Weights for BCA by Combining the Economics of VSL and US Policy." Society for Benefit-Cost Analysis Blog.

Fleurbaey, M. (2011). "Willingness-to-pay and the equivalence approach." Revue d'économie politique 121(1): 35-58.

Hahn, R. W. and P. M. Dudley (2020). "How well does the US Government do benefit-cost analysis?" Review of Environmental Economics and Policy.

Hammitt, J. K. (2015). "Implications of the WTP–WTA disparity for benefit–cost analysis." Journal of Benefit-Cost Analysis 6(1): 207-216.

Hammond, P. J. (1992). Harsanyi's utilitarian theorem: A simpler proof and some ethical connotations. Rational Interaction, Springer**:** 305-319.

Harsanyi, J. C. (1955). "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility." Journal of political economy 63(4): 309-321.

Hendren, N. (2020). "Measuring economic efficiency using inverse-optimum weights." Journal of public Economics 187: 104198.

Herrera-Araujo, D. and L. Rochaix (2020). "Does the Value per Statistical Life vary with age or baseline health? Evidence from a compensating wage study in France." Journal of Environmental Economics and Management 103: 102338.

Hugonnier, J., F. Pelgrin and P. St-Amour (2022). "Valuing Life as an Asset, as a Statistic, and at Gunpoint." The Economic Journal 132(643): 1095-1122.

Kaplow, L. (2005). "The value of a statistical life and the coefficient of relative risk aversion." Journal of Risk and Uncertainty 31(1): 23-34.

Kniesner, T. J., W. K. Viscusi and J. P. Ziliak (2010). "Policy relevant heterogeneity in the value of statistical life: New evidence from panel data quantile regressions." Journal of Risk and Uncertainty 40(1): 15-31.

Layard, R., G. Mayraz and S. Nickell (2008). "The marginal utility of income." Journal of Public Economics 92(8-9): 1846-1857.

Liou, J.-L. (2018). "Effect of Income Heterogeneity on Value of Statistical Life in Taiwan." Available at SSRN 3246540.

McCarthy, D., K. Mikkola and T. Thomas (2020). "Utilitarianism with and without expected utility." Journal of Mathematical Economics 87: 77-113.

Milleron, J.-C. (1972). "Theory of value with public goods: A survey article." Journal of Economic Theory 5(3): 419-477.

Murray, J. S. (2018). "Multiple imputation: a review of practical and theoretical findings." Statistical Science 33(2): 142-159.

Nozick, R. (1974). Anarchy, state, and utopia, new york: Basic Books.

Robinson, A. and A. Spencer (2006). "Exploring challenges to TTO utilities: valuing states worse than dead." Health Economics 15(4): 393-402.

Robinson, L. A. (2020). "Policy MonitorHow US Government Agencies Value Mortality Risk Reductions." Review of Environmental Economics and Policy.

Rowen, D. and J. Brazier (2011). "Health utility measurement."

Sen, A. (1986). "Social choice theory." Handbook of mathematical economics 3: 1073-1181.

Somanathan, E. (2006). "Valuing lives equally: Distributional weights for welfare analysis." Economics Letters 90(1): 122-125.

Torrance, G. W. (1986). "Measurement of health state utilities for economic appraisal: a review." Journal of health economics 5(1): 1-30.

Von Neumann, J. and O. Morgenstern (1947). "Theory of games and economic behavior, 2nd rev."

Weymark, J. A. (1985). "Money-metric utility functions." International Economic Review: 219-232.

Weymark, J. A. (2016). "Social welfare functions." The Oxford handbook of well-being and public policy 1.

On Line Appendix: Compensating variation for mortality risk and the money value of a statistical life

The usual approach to the money value of a statistical life is to consider expected utility given by

$$p\,\mathrm{u}_a(m) + (1-p)\,\mathrm{u}_d(m)$$

where p is the survival probability , m is baseline wealth, and $\mathrm{u}_a(m)$ and $\mathrm{u}_d(m)$ are the utility of wealth if alive and dead respectively and we assume $\mathrm{u}_a(m) > \mathrm{u}_d(m)$. The utility of wealth if dead may be non-zero if there is a bequest motive. Expected utility at baseline is

$$U_0 = p_0\,\mathrm{u}_a(m_0) + (1-p_0)\,\mathrm{u}_d(m_0)$$

where p_0 is the baseline survival probability ,and m_0 is baseline wealth.

Now consider an increase in the survival probability by the amount p but with a compensating variation in wealth of amount wealth CV(p) to keep the person indifferent between the change and the baseline. That is

$$(p_0 + p)\,\mathrm{u}_a(m_0 - CV(p)) + (1 - p_0 - p)\,\mathrm{u}_d(m_0 - CV(p)) = p_0\,\mathrm{u}_a(m_0) + (1-p_0)\,\mathrm{u}_d(m_0)$$

Totally differentiating both sides with respect to p gives

$$\mathrm{u}_a(m_0 - CV(p)) - \mathrm{u}_d(m_0 - CV(p)) - \left[(p_0 + p)\,\mathrm{u}_a{}'(m_0 - CV(p)) + (1 - p_0 - p)\,\mathrm{u}_d{}'(m_0 - CV(p)) \right] CV'(p) = 0$$

We can define the solution of this equation at p=0 as the value of a statistical life; the marginal rate of substitution between survival probability and wealth at the baseline. For small changes in p the compensating variation will be proportional to the MVSL.

MVSL= $$CV'(p)\Big|_{U=U_0} = \frac{\mathrm{u}_a(m_0) - \mathrm{u}_d(m_0)}{p_0\,\mathrm{u}_a{}'(m_0) + (1-p_0)\,\mathrm{u}_d{}'(m_0)}, \quad CV(p) \approx pMVSL$$

To take a simple example let us assume $p_0 = 1, \mathrm{u}_a(m) = \sqrt{m}, \mathrm{u}_d(m) = 0$.

Since we have defined the baseline survival to be 1 we can only consider negative values of p. The compensating variation can be defined by

$$(1+p)\sqrt{m_0 - CV(p)} = \sqrt{m_0}$$
$$CV(\mathrm{p}) = m_0 \frac{(1+p)^2 - 1}{(1+p)^2}$$

Since we have taken the baseline to be a survival probability of 1, the changes p have to be negative and V(P) the compensating variation is also negative; money has to be paid to have the agent accept a risk of death

Now let s be the total survival probability s=1+p after the change p then we have the payment needed to make the person accept this survival probability rather than stay as the baseline survival 1 is minus the compensating variation :

$$CV(s) = m_0 \frac{1-s^2}{s^2}$$

We can also calulate the equivalent variation EV(p)

$$(1+p)\sqrt{m_0} = \sqrt{m_0 + EV(p)}$$
$$EV(\mathrm{p}) = ((1+p)^2 - 1)m_0$$

The money the person is willing to pay to stay at alive with certainty rather than undergo a survival risk s is
$$EV(\mathrm{s}) = (1-s^2)m_0$$

Note that willingness to pay to avoid a risk of death is bounded above by wealth while the compensation needed to undergo the risk may be unbounded.

While the MVSL and approximation to the compensating variation is
MVSL= $V'(p)\big|_{p_0, m_0} = 2m_0 = 2, \quad V(p) \approx 2m_0 p = 2m_0(1-s)$

Figure A1 shows the compensating variation CV, the equivalent variation EV, and the linear approximation using the MVSL for different total survival probabilities, measured in multiples of baseline wealth. For survival probabilities close to one, the MVSL approximation is close to both the compensating and equivalent variation. The equivalent variation rises more slowly than the linear approximation and always lies below total wealth. However, the compensating variation is highly non-linear and is much higher than the MVSL approximation for low survival probabilities.

Figure 1A depends on our choice of the square root utility function, different functions will be have a more or less non-linear compensating variation. We find very large payments being required to compensate people for a high risk of death ex ante. Note that this is different from the issue of compensating the dependents of people who actually die ex post in legal liability cases where the ex post perspective makes the use of ex ante preferences unappealing (Becker and Stout 1992).

Figure 1A makes to points. The first is that the MVSL is only a good approximation for small mortality probabilities; one the probability of death becomes large is may not be a good guide to willingness to pay. The second is that the large gap between an individual's EV and CV for high probabilities of death is likely to make money metric cost benefit analysis inconclusive. Given that the compensating variation needed to accept death with certainty is unbounded, no policy choice that implies certain death for an individual can ever be desirable from a money metric cost benefit perspective. Similar inconclusive results are likely when the risk of death is non-marginal, and EV the CV diverge.

Again returning to our problem of choosing between a outcome with 100 million people dying and an alternative with 1 different person dying the unbounded compensating variation of the one person makes the choice inconclusive for cost benefit analysis. However, the problem persists even for lower but non-marginal probabilities of death. Suppose we have 4 people with same wealth and have to choose between a policy that imposes a probability of death of 0.5 on three of them or an alternative policy that imposes a probability of death of 0.5 on the other. In our example the compensating variation needed to undergo a probability of death of 0.5 is 3 times wealth while the equivalent variation is 0.75 times wealth. This means that the 3 would be willing to pay only an aggregate of 2.25 times their initial wealth levels, less than required to compensate the one. Clearly the one cannot compensate the 3 for undergoing the alterative where they take the mortality risk, and money metric cost benefit analysis unable to rank the two alternatives. Only when the number involved exceeds 4, and they can compensate the one, will cost benefit analysis prefer one person having the mortality risk to many having the risk.

While the money metric compensating and equivalent variations are non-linear in the probability of death, and may be unbounded, our expected utility approach is linear in probabilities. Taking the utility of death to be zero and the utility of being alive to be one the expected utility approach will always lead to minimizing the expected number of deaths when deciding between outcomes with different distributions of mortality risk.

Figure 1A: Compensating Variation, Equivalant Variation, and MVSL Approximation by Survival Probaility