Preference Calibration with QALYs

V. Kerry Smith[*]
North Carolina State University
kerry_smith@ncsu.edu

Subhrendu K. Pattanayak[**]
Research Triangle Institute
subrendu@rti.org

George L. Van Houtven[**]
Research Triangle Institute
gvh@rti.org

March 2003

[*] University Distinguished Professor, Department of Agricultural and Resource Economics, North Carolina State University and Resources for the Future University Fellow.
**Senior Economist, Research Triangle Institute and Visiting Assistant Professor, North Carolina State University; and Senior Economist, Research Triangle Institute, respectively.

Support for this research was provided by EPA Cooperative Agreement CR824861-01-0. Smith's work on this paper was supported by subcontract #17-42U-6600 from RTI International

"Money can't buy happiness."

1. Introduction

This homespun advice serves as a tangible reminder that some essential sources for utility are derived outside markets. Individual health is one of them. Environmental quality is another. Economic analyses of the effects of these non-market sources of utility for individual behavior usually highlight the extent to which they have private good attributes. These discussions then turn to people's ability to adjust their consumption in the absence of markets. Such substitution opportunities have important implications for situations where it is argued there is a policy role in changing the amounts of one or more non-market commodities.

Health and environmental policies are prominent examples that fit this description. In both cases one finds concern raised with the use of monetary measures of the benefits provided by these programs. Non-economists routinely seek some alternative index to gauge the "output" of policy. In areas where health and environmental policy affect the same outcome, they do not always share the same output measure. With fatality risks they do, while for morbidity effects they do not. In the latter set of health effects, environmental policy analyses have struggled with developing measures that could be used in assessing monetary benefits of avoiding specific types of negative health impacts, while analyses of health policy have focused on quantity indexes that attempt to take account of severity and duration of health impacts in a single index allowing aggregation across people. Quality adjusted life years (QALYs) are measured considering both the duration of a health condition and the health-related quality of life associated with that condition.

Recently, there has been interest in evaluating whether the assumptions made in defining QALYs would allow consistent monetary measures of the benefits of increased QALYs (Hammitt [2000b]). The framework used to define the QALY index does not readily conform to most of the preference specifications used to describe important economic choices. As Hammitt concludes:

> "If an individual's utility function for health and longevity can be represented by the number of risk-adjusted QALYs, and if his tradeoff between health and longevity is independent of wealth (as routinely assumed) then his utility function for health, longevity and wealth is tightly constrained" (p.15).

Rather than work within this constrained setting we propose an alternative strategy. The QALY elicitation format is one that appears to have been effective in eliciting health choices to gauge the importance of their severity and duration. Given this premise, we propose that the choices underlying QALY measures be interpreted using conventional preference specifications. These interpretations allow the results from QALY elicitations to be used along with other economic information about behavior to calibrate the health related components of individual preference functions. This practice assures any index of the health condition derived from the calibrated preference function will be consistent with economic benefit measures.

To develop some background for our approach, section two discusses potential explanations for concern with economic measures of outcomes and the factors influencing heterogeneity in marginal values for non-traded goods across people. This discussion is intended to provide one motivation for using QALYs. The section closes with a discussion of how QALYs are defined and measured. In section three we outline the structure of preference calibration, review how it was used with job risk information, and then demonstrate how QALY information could be used in preference calibration. Our strategy develops the logic in a sequence of progressively more complex examples. Section four uses the intermediate case to present some illustrative preference calibrations and the last section highlights some issues that should be considered in next steps in this research.

2. QALYs and Economic Tradeoffs

A quality adjusted life year (QALY) is an index derived by eliciting an individual's evaluation of a health state. It combines the subjective evaluation of the health condition along with its duration. When there is one health state that lasts for a time span, T, then the QALY index is defined as a simple product – $q \cdot T$, with q a numerical gauge of the quality of the health index, using a scale between zero and one. A value of one for q designates perfect health and zero death.[1]

QALYs are often derived by presenting individuals with economic tradeoffs. One common approach designated as the standard gamble offers a person a choice of some undesirable health condition with a specified duration in comparison to a lottery of perfect health for that duration versus death. Each individual is asked about the probability of death that would be considered equivalent to the undesirable health state. Another elicitation scheme, labeled the time tradeoff, asks about years of longevity in perfect health a person would give up to avoid an undesirable health state. Each of these

[1] Sometimes values less than zero are used to describe states perceived by individuals to be worse than death.

strategies compares a situation with one health state to another situation with a different health condition. The comparison is very similar to the types of tradeoffs used to define economic values. In this section we will compare a QALY tradeoff to the ones used to define economic values. Sometimes the health state resembles one the respondent has experienced. In other cases both are hypothetical situations. The analysis of the responses to QALY tradeoffs is different from the analysis of tradeoffs to measure structural economic relationships. More specifically, the QALY framework begins with an assumption that the index takes a particular form, can be aggregated across health conditions, durations, and people. Given those assumptions, the analysis usually asks what preference specifications will be consistent with using the answers these ways. By contrast, economic models take the preference specification as the fundamental starting point and the properties of the index follow from it.

Several aspects of the comparison are relevant to efforts to integrate QALYs into a purely economic framework for evaluating outcomes. We begin this process by discussing some areas of concern with economic values, then turn to a discussion of how uncertainty affects the definition of economic values, and close this section by describing how the definition of the QALY health index can be interpreted as one approach for addressing the concerns about choices involving health states and uncertainty. Our overall conclusion is that QALYs do not resolve the problems with economic values. As a result, we believe there is a need to consider more specific ways to integrate the health index offered by QALYs with more conventional economic framework. We propose to extend preference calibration as one strategy for meeting this objective.

2.1 Economic Tradeoffs under Certainty

An economic value is defined by describing a tradeoff. That is, some change in a good or service is described by evaluating two alternative choices – more of one thing and less of another. These comparisons define economic values in relative terms. In formal terms an economic value is defined by a marginal rate of substitution – the rate of exchange between two commodities that would not change an individual's level of utility. In the thought experiment used to define an economic tradeoff when an individual receives more leisure time in a week, he (or she) must give up something else to assure well being is unchanged. For example, we might compare the added time to discretionary resources for clothing. This type of comparison is an analytical one. We do not, as a rule, observe the exact tradeoffs that define these points of indifference. Nonetheless, for private goods with ideal markets, we can indirectly measure these marginal values. That is, we assume people's choices arise from budget constrained utility maximization. This process implies for those private goods available in markets that relative prices will reveal these marginal values.

It also implies something else that can be easily overlooked. Market exchange assures that all people adjust until their marginal values are equalized. We count on this outcome in the argument that holds ideal markets assure an efficient allocation of private goods. By equalizing the marginal values of private goods across all individuals we can be sure there is no further reallocation of these goods among people that would improve anyone's well being.

This logic implies that if we want to evaluate an exogenous marginal increase in any one of these private goods, all people in the economy will value that change the same way. Moreover, their individual marginal values will correspond to each good's price relative to some numeraire. This relevation property has been very important to evaluations of programs that provide private goods. The analyst can assume regardless of who received the services we know the right weights to use in counting up the "worth" of the outputs provided to a diverse array of people. The dollar weight assigned to additional amounts received by any individual are the same. People seem to be treated equally. This outcome follows from the special circumstances defining the goods and their availability independent of the exogenous public program being evaluated. Once we leave the domain of marketed private goods to examples where the goods are private, but not freely traded, there is no longer assurance that the marginal values will be equalized.

Consider the case of quasi-fixed private goods. Incremental values or virtual prices may well be quite different across consumers. If we treat an individual's health as a private commodity, then we should expect the amounts available will be different and there is not a market that allows people to "trade" health services. At the margin, each person can be expected to have different values for small changes in his or her health. This interpretation of health as a quasi-fixed commodity may help to make it clearer why these discrepancies in values across people become of interest. Any population can in principle have very different values for programs to improve individual health.

Personal health as a commodity cannot be traded among people or across time, so there may be large differences in the incremental values specific people will place on changes in this commodity. To establish this point we need only write the indirect utility function corresponding to budget constrained utility maximization and define the

Marshallian virtual prices. Equation (1) provides such a general definition with *m* the income, *p* a vector of prices of marketed goods, and z_1 and z_2 two quasi-fixed commodities.

$$(1)\quad V = V(p, m, z_1, z_2)$$

The virtual price (Marshallian) for z_1, π_{z_1}, is defined in equation (2).

$$(2)\quad \pi_{z_1} = \frac{V_{z_1}}{V_m} \qquad \text{with } V_j = \frac{\partial V}{\partial j}$$

There is no mechanism that assures different people will assign the same incremental value to z_1. The diversity in marginal values across people for private goods that do not exchange on markets offers one explanation for why policy analysts seek a standardized index to "keep score" of the incremental value of adding to the z_1 or the z_2.

Most discussions of alternatives to economic values for health (and environmental resources) do not cite this issue of diversity in marginal values across people.[2] They tend to focus instead on the absence of substitutes, the inability to trade through time, and uncertainty. The potential for very different values across people with no opportunity for arbitrage to reduce the discrepancy also seems another reason for greater attention to how a numeraire or reference point affects the importance that is attached to particular types of tradeoffs.

Of course, it is also possible to have situations outside formal markets, involving marketed goods that can serve to reduce the diversity in incremental values for non-traded goods. For example, substitutes for these goods can mitigate the diversity in

[2] Brekke's [1997] discussion of the influence of the numeraire is a notable exception. He discusses how the numeraire can influence judgments about the merits of programs providing public goods when individuals' preferences are heterogeneous in their evaluations of non-traded goods.

marginal values across people. To establish how this process would work, consider the case of perfect substitution. That is, a case where the non-marketed private good has a market good which is a perfect substitute. The non-market good need not be a public good. In this case, it is the market prices of this other good together with the high level of substitution that brings marginal values into alignment across people.

The market prices for the marketed private good serve as the basis for incremental values. To illustrate, let x_1 be a private good with price p_1, serving as a perfect substitute for z_1 at a fixed rate, c, (i.e. there is a sub-function in preferences linking the two goods, $q = x_1 + c \cdot z_1$). Then the disparity in incremental values across people is "adjusted away" by their actions (assuming c is the same). This point is illustrated by making some substitutions outlined in equations (3a) through (3c).

(3a) $-m_{z_1} = p_1 \cdot c$

(3b) $V_{z_1} = -V_m m_{z_1}$

(3c) $\pi_{z_1} = \dfrac{V_{z_1}}{V_m} = \dfrac{V_m m_{z_1}}{V_m} = p_1 \cdot c$

Most private goods we can substitute over time. If preferences are additively separable and the individual rate of time preference is the same as the market rate of interest, then the relative marginal values of private goods will be adequately captured by relative prices and the conclusion that market forces lead to an equalizing of the marginal values across people taking account of timing will be upheld.

The same is not true for health. Not only is it impossible to substitute health conditions over time, but there are also important elements of irreversibility present in many types of changes in health conditions. Serious changes to an individual's health status – arthritis, heart disease, and other degenerated diseases – may not be an immediate threat to an individual's life, but may lead to irreversible changes in health status. Thus, for private goods that do not exchange on markets we face an issue that is comparable to public goods – the potential for significant diversity in the incremental values of these commodities across people and little opportunity to completely adjust.

Often these commodities, or the opportunities to adjust to them, involve people in choices in the presence of uncertainty. Thus, the relevant incremental values must include recognition of differences in attitudes toward and abilities to adjust to risk.

2.2 Tradeoffs with Uncertainty

We will use a state claims framework to describe tradeoffs under uncertainty. By adopting a state dependent specification for preferences, this model acknowledges that there are some situations involving risk, such as a health outcome, where it is not possible to describe in a simple algebraic form, as an index, how an individual's preferences will change with the event at risk.[3] We do acknowledge at a general level that the marginal utility of income will be different across states and this is the key feature distinguishing how tradeoffs are represented. Three features are important to tradeoffs under uncertainty: (1) the distinction in the standard used to measure individual well being; (2) the monetary measure used to describe economic value; and (3) the opportunities assumed to be present for private adjustment.

[3] As we move from the general conceptual logic to preference calibration discussed in the next section, the analysis must become explicit and, as a result, the generality of the basic framework can not be maintained. We must specify a form for preferences in order to describe how choice information links to specific features of the maintained preferences.

To use the same general notation as in the previous section, but focus on the key issues relevant to uncertainty, we begin by dropping the quasi-fixed arguments of $V(\cdot)$ and add a subscript to identify utility as dependent on the state of nature – A (good health) and D (bad health or death). In keeping with the generality often used when this framework is used to explain job risk choices, we do not explain at this stage how the events affect utility and instead acknowledge that they do with the subscript. P is the probability of the bad health state. Expected utility is usually assumed to describe the individual's objective function as in equation (4).

$$(4)\ EU = (1-P)V_A(p,m) + PV_D(p,m)$$

In this general statement we have not allowed for state dependent adjustments such as insurance. These opportunities can be introduced by allowing state dependent choices and, with them, different prices. We will consider how this added opportunity affects the analysis of ex ante monetary measures for willingness to pay below.

The ex ante marginal rate of substitution used to define the value of a statistical life (VSL) is readily defined by assuming the state B is death and assuming the individual selects among jobs that vary in compensation for accepting the risk of death, P, through a wage function, $W(P)$. With a continuous array of choices, the incremental compensation will be equal to the ex ante marginal rate of substitution as represented in equation (5).[4]

$$(5)\ \frac{dW}{dP} = -\frac{EU_P}{EU_W} = \frac{V_A - V_D}{(1-P)V_{Am} + PV_{Dm}}$$

The terms V_{Am} and V_{Dm} designate the marginal utility of income in each state. Often the analysis of VSL behavior assumes that the level of utility in the state D, V_D, is not related

[4] To derive this result, we assumed that all of the individual's income was derived from wages. This implies that m is replaced by $W(P)$ in V_A and V_D.

to the available income, *W(P)*. In the extreme, if the state was death then the income defined as wages would not be received if an individual was dead. Thus, in this case only the marginal utility of income in the state good health contributes to VSL and equation (5) would be rewritten as equation (5a).

$$\text{(5a)} \quad \frac{dW}{dP} = \frac{V_A - V_D}{(1-P)V_{Am}}$$

To appreciate how this framework relates to a range of economic tradeoffs and QALYs it is helpful to use Cook and Graham's [1977] early discussion of how state dependent preferences influence the way we characterize the importance of the events at risk, the effectiveness of insurance, and the ex ante value of policies to reduce risk. These authors introduced a direct measure of the uniqueness of states by defining the willingness to pay for a certainty of the more desirable state as compared to the less desirable one, as in equation (6).

In some situations we would be comparing states other than life and death. A person might compare life with near perfect health to a situation with a chronic illness. In this case, $\theta(m)$ would gauge the relative importance of protecting perfect health over that alternative.

$$\text{(6)} \quad V_A(p, m - \theta(m)) = V_D(p, m)$$

$\frac{d\theta(m)}{dm}$ can be used to describe whether the conditions characterizing the two states are unique or replaceable by measuring the disparity in the marginal utilities of income between the two states when their total utilities are held equal through the construction of the point of comparison.

(7) $$\frac{d\theta(m)}{dm} = 1 - \frac{V_{Dm}(p,m)}{V_{Am}(p,m-\theta(m))}$$

Figure 1 illustrates the concept graphically. In this case, the indifference curves correspond to the probability weighted sum of V_A and V_D. As a result, the set of ex ante indifference curves change because of changes in P. This change is analogous to what changes in a third good might do to an indifference map.[5] The layout of the graph allows for state dependent income payments. It is a common format used to illustrate how people select insurance with different opportunities to adjust to alternative events at risk. For this analysis, our primary goal is to use the framework to describe how the VSL, option price, ex ante willingness to pay, and QALYs relate to each other.

We begin the process noting that separate labels, m_A and m_D, imply it is possible to allocate income between the two states. If we assume the "prices" of insurance correspond to actuarially fair rates then we can describe how insurance affects behavior and with it the economic values for risk changes with different opportunities to adjust privately to the events at risk.[6] The ex ante insurance constraint [e.g. $m = (1-P)m_A + P \cdot m_D$] describes the "cost" of having state dependent income. Our bracketed definition implies the prices correspond to the probabilities for each state. Tangency of insurance "budget" constraint with the ex ante indifference curve illustrates how access to this opportunity to adjust to risk influences the reallocation of resources

---

[5] In fact, this shifting of indifference curves to reflect constant utility but different levels of a third good was a strategy used to illustrate complementarity between goods. See Smith and Banzhaf [2003] for an example.

[6] The first order conditions for optimal insurance equate the ex ante marginal rate of substitution to the relative prices of state dependent claims. When the relative prices of these claims equal relative odds, it is labeled actuarially fair.

$$\frac{EU_{m_A}}{EU_{m_D}} = \frac{(1-P)}{P} = \frac{(1-P)V_{Am}}{PV_{Dm}}$$

between states. By including the certainty locus on the graph, it is also possible to describe how allowing state dependent income allocations permits an evaluation of the importance of the uniqueness of the situations that define the state dependent preferences.

To develop this comparison we use several aspects of the diagram. First, by selecting a value of income, *m*, on the horizontal axis, it is possible to illustrate the certainty locus. This locus defines the amount of income in state *A*, a person would give up to avoid the state *D*. Thus, at *m*, $\theta(m)$ is measured by the difference between the certainty locus and the 45° line (i.e. $V_A(m-\theta(m)) = V_D(m)$). It is identified on the vertical axis. The gauge of irreplaceability of specific states or goods (measured by how $\theta(m)$ changes with *m*) refers to a comparison of the value of an added dollar of income in each state when an individual's total utility in each state is constrained through $\theta(m)$. In Figure 1, one way of evaluating this point considers the tangency of the insurance budget constraint, drawn here for the case of fair insurance, with the ex ante indifference curve. With fair insurance this tangency arises where the marginal values of added income are equal across states. Nonetheless, at this income level in the state *A* (m_A^*) an individual would need to give up income to be indifferent (in terms of total utility) with the state *D*. Fair insurance does not offer enough "protection" against the loss in well being attributed to the state *D*. This point is illustrated by considering the intersection of a vertical line at m_D^* with the certainty locus. The income level at *A* yielding the same level of total utility in *A* is substantially less than m_A^*. Thus, even if market adjustment allows the marginal value of income allocated between the two states (in ex ante terms) to be equalized, a person would still prefer *A*.

When these states correspond to alive and dead, this result is hardly surprising. However, if the states were alive and another non-fatal health condition, measuring how θ behaves with income allows a judgment about whether ideal adjustment to risk fully mitigates the conditions a person could experience. In the example given in figure 1 it does not. Ideal insurance would not be sufficient compensation. These distinctions affect different ex ante measures of economic value in different ways.

Figure 2 this Cook-Graham graph to illustrate further the differences between two measures and equations (7) and (8) describe in more specific terms how the good's uniqueness and the opportunities for adjustment influence the economic value of risk changes.

Consider first the example given in figure 2. Suppose that the probability of the undesirable state (*D*) declines by ΔP. This change leads to a shift out in the ex ante indifference locus from $EU(P)$ to $EU(P-\Delta P)$. Maintaining fair insurance leads to a non-parallel shift in the insurance opportunities locus, flattening the locus. The optimal mix of insurance changes from I to II. We can compare the economic value of that risk change by measuring along the 45˚ line. The difference between the intersections of the two insurance loci (or ex ante budget constraints), *CD*, measures the incremental value of the probability reduction.

By contrast, we might consider, as the VSL implicitly does, a state independent payment for the probability reduction. This value concept would be an option price. It requires the largest ex ante payment, constant across states an individual would commit to realize the probability reduction. The intersection of $EU(P)$ and $EU(P-\Delta P)$ with the

45° line allows this measure to be compared with the value assuming fair insurance. It is given by *AB*. In general, *CD* will exceed *AB* (Graham [1981]).

If we define *m(·)* as the ex ante expenditure function (i.e. the total expenditure in the two states required to hold expected utility constant), then we can describe the value of a probability change with fair insurance in equation (7) and the incremental option price in (8). Fair insurance simplifies the denominator of the second term in equation (7) because marginal utilities are equalized across states through the insurance market. With the option price m_A^* is constrained to equal m_D^* so the first term drops out.

$$(7)\ \frac{\partial m}{\partial P} = m_A^* - m_D^* + \frac{V_A(m_A^*) - V_D(m_D^*)}{V_{Am}} \qquad \text{irreplaceable (with fair insurance)}$$

$$(8)\ \frac{\partial m}{\partial P} = \frac{V_A^* - V_D^*}{(1-P)V_{Am} + PV_{Dm}} \qquad \text{equal payments (option price)}$$

If we consider a case where the conditions associated with *A* and *D* are "replaceable" (that is, total utilities are equal under fair insurance $V_A(m_A^*) = V_D(m_D^*)$) then the second term drops out. Fair insurance would measure the gain (or loss) from probability changes as indicated with equation (9).

$$(9)\ \frac{\partial m}{\partial P} = m_A^* - m_D^* \qquad \text{replaceable (with fair insurance)}$$

Self insurance and mitigation can be discussed (see Shogren and Crocker [1991, 1999]) as complicating the predictions that can be made about the relationship between probability changes and behavior.[7] For our purposes they are analogous to the case of

[7] These cases allow for private, non-market behaviors that are analogous to market insurance and private actions that make the *P* endogenous for each individual.

quasi-fixed goods in that they offer another non-market avenue for differences in individual values for risk.

Our analysis to this point has indicated that once uncertainty and state dependent preferences are allowed, there are added reasons to expect diversity in the incremental values of states of nature that can be considered analogous to private goods and, in this case, are associated with risks. To the extent these risks can be "chosen" in a market process, as would be the case in job related risks, and market based compensation is also associated with those choices, then each person adjusts his or her circumstances to a market schedule. This process does not imply equalization of marginal values across people as is the case with fixed market prices described in the case of certainty. Thus, uncertainty together with non-traded private goods offer potential explanations for the interest in defining another index or numeraire for comparing health outcomes across people. The objective appears to be one that defines a set of weights so it is possible to treat comparable health conditions and the durations (over which these conditions are experienced) in the same terms. As we discuss in the next section, these features are assumed to characterize the QALY index. In practice, with nearly all plausible specifications of preferences, they would not be satisfied by the indexes derived using standard gamble or time tradeoff methods (see Hammitt [2000b]). That is, the properties attributed to QALYs would not be consistent with most conventional specifications of preferences.

2.3 QALY as an Economic Index

Broome [1993] and Hammitt [2002a, 2002b] among others have written extensively on the economic foundations of the QALY scale. Following Pliskin et al. [1980], Hammitt summarizes the four assumptions required of individual preferences for QALYs to serve as a consistent index of individual preferences. They are:

A. *Mutual Utility Independence*

   This requires preference between lotteries for health states to be independent of life span and preferences between lotteries for life span to be independent of health state.

B. *Constant Proportional Tradeoff of Longevity for Health*

   The fraction of an individual's remaining life span a person would trade to improve health is independent of his or her remaining life span.

C. *Risk Neutrality Over Life Span*

   Holding health state constant an individual prefers lotteries with longer to shorter expected values for life expectancy.

D. *Additive Independence Over Time*

   An individual's preferences for lotteries on health in any subset of life span do not depend on health in other times.

These conditions are extremely demanding. Broome's summary of the properties of utility functions required to meet the ways the QALY index is used seems to capture these assumptions in a direct way. He suggests that preferences conform to expected utility and that they be consistent with risk neutrality. Finally, the QALY properties require that individual preferences must be additive in QALYs (or discounted QALYs).

These features imply when we consider either the standard gamble or time tradeoff elicitation questions there is a direct link between the response given by an individual and the health or time component of the score used to create the QALY index. Consider three examples – two with a standard gamble and one with time tradeoff. For the first, suppose an individual is confronted with a certainty of health condition $\overline{H}$ for T periods versus a lottery with a probability π of excellent health (H^*) for T periods and ($1-\pi$) of death. When an individual is asked what π would make him (or her) indifferent between the health conditions and the lottery, we are asking for a probability defined through an expression for a certainty equivalent. In Broome's framework this amounts to solving equation (10) for $g(\overline{H})$, where $V(Dg(\cdot))$ is the preference function, and D is the discount factor (either T with no discounting or $\sum_{i=1}^{T}\gamma_i$ with γ_i the discount factor in the ith period).

$$(10)\quad V\left(Dg\left(\overline{H}\right)\right)=\pi V\left(Dg\left(H^*\right)\right)+\left(1-\pi\right)V\left(0\right)$$

Assuming $V(0) = 0$, because V is taken to be linear and $g\left(H^*\right)$ is normalized to unity we have $g\left(\overline{H}\right)$ revealed through the responses given for π. If the lottery were expressed in terms of time, it would hold π constant in the state $\overline{H}$ and H^* and vary T, say $T-\Delta T$ in H^* versus T in $\overline{H}$, with some normalizations and no discounting $g\left(\overline{H}\right)=\left(T-\Delta T\right)/T$. Finally, if we consider a straight time tradeoff and no comparison of lotteries we realize the same outcome as with the lottery, due to the assumption of linearity and the normalization of $V(0)$.

Thus, the ability to treat QALYs as constant across people (and time) and the "conclusion" that it is revealed by standard gambles or time tradeoff results from the assumptions about preferences. It is not an inherent feature of the QALY concept. Indeed, we would expect QALYs would be subject to considerable heterogeneity, reflecting the non-market attributes of health as a commodity. This point is simply a restatement of our argument that when people do not have the ability to adjust levels for private, rationed commodities, they will attach different levels of importance to marginal adjustments in these rational goods (i.e. they have different virtual prices). These differences can be expected to be reflected in how they respond to the standard gamble or time tradeoff questions. Imposing the QALY structure of preferences on the interpretation we derive from their responses simply selects a format for approximating preferences at a designated point. Its properties will depend on the nature of the approximation error.

Hammitt's [2002b] recent analysis of admissible functions using risk adjusted, quality adjusted life years implies that an individual's willingness to pay per QALY will not be constant and can be expected to depend on health, longevity, wealth and attitude toward risk with respect to longevity. In short, the consistency across different people is lost.

We propose an alternative interpretation, which considers another way to use the responses to QALY elicitations. We propose to use those responses in a way that is broadly consistent with the QALY treatment of the health condition and its duration, but allows the responses to be more directly included within conventional preference specifications. Thus, rather than seeking a function that meets QALY restrictions with respect to economic variables like wealth or income, we interpret QALY responses using a model that is derived from a well-defined preference function. It implies the issues in reconciling QALYs with economic measures of WTP are not ones that require preference specifications where both can be consistently defined. Instead, we view both frameworks as sources of choice data that can be used to calibrate a common function.

3. Preference Calibration and QALYs

3.1 Background

Preference calibration is motivated by the need to use past empirical estimates describing revealed or stated choices to infer the benefits from new policies. These interventions involve changes in resources that may not be represented by a set of empirical estimates in the available literature. As a result, there has been a wide range of research investigating whether normalized measures of the consumer surplus or willingness to pay, standardized by some measure of the amount of the resource change considered in each study, offer a reasonable basis for transfer. These measures have been used as constants or with adjustments developed from statistical models that attempt to control for features of the existing research.[8]

By contrast, calibration imposes the consistency of a preference function on the ways existing information is used in a new situation. It requires the specification of a preference function. This is usually an indirect utility function that is specified in a format consistent with the sources of available data.[9] The preference function is treated as describing the representative individual's utility and the task of transfer is rephrased as an identification problem. It asks whether the available benefit estimates and

[8] Mrozek and Taylor's [2002] meta analysis including variables that account for features of the specifications and data used to estimate wage hedonic models would be an example. In their case they use the statistical summary to propose an adjusted VSL.

[9] This consistency amounts to including variables that would relate to the benefit measures available. If the available studies are primarily recreation, then the indirect utility function would be expressed in terms of travel costs, income, perhaps wages, and site qualities. If they were from hedonic property values then indirect utility functions consistent with these models would be the focus.

supplementary information about the individuals involved allow analysts to identify estimates of those preference parameters required to define a consistent valuation function (e.g. a Hurwicz-Uzawa income compensation function, see Hanemann [1999] for discussion).

We have outlined how these methods could be used to combine travel cost, contingent valuation, and hedonic property value estimates for recovering valuation functions for water quality. We have developed analogous structures describing how wage hedonic estimates of the ex ante value of risk reductions, together with contingent valuation measures of the option price for risk reductions and labor supply models, might be combined to identify the parameters of preference functions and used to calibrate option price functions for valuing reductions in fatality risks. Here we propose to extend this second line of research to include information from QALY elicitations using the standard gamble format to calibrate a health state component of a representative individual's preferences.

Because our analysis applies to a single, representative individual, this strategy offers the opportunity to assure greater consistency in the link between QALYs and WTP for policies that would lead to changes in health conditions. It also permits greater uniformity in treating different individuals the same way in so far as evaluating the importance of health gains or losses. These objectives are met by defining reference conditions that underlie the health index function and then translating the standardizing changes into the implied variations in health indexes for different individuals so the associated economic valuation measures can be consistently defined for the relevant "reference commodity".

This strategy contrasts with the discussion in the previous section in that it requires a specific algebraic function be identified to describe preferences for the representative individual. The logic underlying this formulation is not based on early arguments associated with representing aggregate responses as if they could be captured by a single individual's preferences (see Deaton and Muellbauer [1980] pp.149-155), but instead as a convenient method to impose preference consistency on benefit transfers. That is, we acknowledge that benefit measures have direct relationships to income and the other features contributing to the constraints associated with people's choices. As we adjust benefit measures to fit the new circumstances associated with different policy applications, there should be consistency with the realization of these types of constraints in the new application. It would be unreasonable to expect people will be able to pay more for a change in some exogeneously determined commodity, health state, or risk than their available income.

Moreover, the amount each person might be expected to be willing to pay would depend on available substitutes, or in the case of risk, the opportunities to adjust or mitigate. The conditions that confront the individuals studied in any particular analysis that is the primary research study used to derive benefit measure need not correspond to those that would confront different individuals (or even the same individuals) in the context that corresponds to a particular policy application. Thus, there is a need for some approach to adjust or adapt benefit estimates to reflect these differences.

3.2 Calibration with Risk Changes

Our application (Smith et al. [2002]) to risk assumed that the analysis included measures of the ex ante marginal rate of substitution for fatality risks from wage hedonics, measures of option price for risk reductions from contingent valuation studies, and information about labor supply choices. In this analysis we followed Burtless and Hausman [1978] and derived an indirect utility function that would be consistent with a simple labor supply model. This specification was then used to define the VSL and option price models.

More specifically, with a labor supply model given in equation (11), the corresponding quasi-indirect utility function is given by equation (12).

$$(11)\quad ln(h) = \alpha + \beta W + \mu m$$

$$(12)\quad V_A = \frac{-exp(-\mu m)}{\mu} + \frac{exp(\alpha + \beta W)}{\beta}$$

where m = non-wage income
W = wage rate
h = hours worked[10]

We labeled the resulting indirect utility as V_A to correspond with our earlier notation.

To complete our description of expected utility and to link the resulting model to the VSL we need to consider the utility for the state death. For this part of the model, it is important to recognize that our description adopts an ex ante perspective. Thus, a utility for the state "death" should be interpreted as an evaluation of that state by an individual at the time the job alternatives and results are "potential" outcomes. From this orientation, $V_D(\cdot)$ likely reflects bequest motives. The function would not include the wage rate, but could be related to available non-wage income. Using V_D as the first term in the right side of equation (13) (dropping the contribution arising from labor / leisure choices) then the expected utility function is simplified. Equation (13) defines the resulting expected utility.

$$(13)\ \ EU = (1-P)\left[\frac{-exp(-\mu m)}{\mu} + \frac{exp(\alpha+\beta W)}{\beta}\right] + P\left[\frac{-exp(-\mu m)}{\mu}\right]$$

We can use this expression to define three relationships. Each corresponds to a connection between what is measured through revealed or stated preference methods and the parameters of the expected utility function assumed to motivate the choice. For example, from hedonic wage models, we have estimates of the ex ante marginal rate of substitution between wages and the risk of death, *P*, on the job. Treating *P* as that risk we can express this VSL in equation (14).

$$(14)\ \ VSL = -\frac{EU_P}{EU_W} = \frac{dW}{dP} = \frac{1}{(1-P)\beta}$$

To define another way in which the available measures relate to this preference function we considered the contingent valuation questions asked in Gegax et al. [1991] because the risks were described as arising from workplace activities. We interpreted their question as asking individuals to respond with an option price – a commitment that would be made out of income before they knew the outcome.[11] Equation (15) translates

[10] The units used to measure hours correspond to what is relevant for the choice model. To link to VSL estimates, annual hours would be relevant.

[11] The Gegax et al. questions were developed in the context of job risks, but asked about annual gross income (without considering a labor supply adjustment). As a result, we interpret them as an option price and define it in terms of *m*. *W* in our model is a rate of pay. Their specific question is given as follows. After presenting a risk ladder and asking each respondent to select a "rung" on the ladder that comes closest to describing the risk of accidental death in their job, the WTA question asks for those below the highest risk:

> "consider a situation in which you were asked to face more risk on your job. What is the smallest increase in annual gross (i.e. before deductions and taxes) income from your job that you would have to be paid in order to accept an increase in the risk of accidental death by one step (i.e. one more death per

the comparison between two job risks, the initial level of P_0 and the new lower level of P, into a statement describing the maximum option price they would pay for the risk change using our hypothesized preference function. P_0 is the initial job risk and P_1 the proposed lower risk that an individual is hypothesized to value at the option price, labeled here as $W\tilde{T}P$.

$$(15)\quad (1-P_0)\left[\frac{exp(\alpha+\beta W)}{\beta}-\frac{exp(-\mu m)}{\mu}\right]+P_0\left[\frac{-exp(-\mu m)}{\mu}\right] = (1-P_1)\left[\frac{exp(\alpha+\beta W)}{\beta}-\frac{exp(-\mu(m-W\tilde{T}P))}{\mu}\right]+P_1\left[\frac{-exp(-\mu(m-W\tilde{T}P))}{\mu}\right]$$

Rearranging terms we can define the option price in equation (16).

$$(16)\quad W\tilde{T}P=\frac{1}{\mu}ln\left[1+\frac{\mu}{\beta}exp(\alpha+\beta W+\mu m)(P_0-P_1)\right]$$

Thus, we have three relations in (17a) through (17c) derived from different types of choices individuals can be observed making in revealed (or stated) preference contexts.

$$(17a)\quad ln(h)=\alpha+\beta W+\mu m$$

$$(17b)\quad VSL=\frac{1}{(1-P)\beta}$$

$$(17c)\quad W\tilde{T}P=\frac{1}{\mu}ln\left[1+\frac{\mu}{\beta}exp(\alpha+\beta W+\mu m)(P_0-P_1)\right]$$

The logic of preference calibration relies on observing measures for *h*, *W*, *m*, *VSL*, *P*, P_0, P_1, and $W\tilde{T}P$ and then solving these three equations for the three unknowns – the parameters (α, β, and μ) required to describe the valuation function for a risk change.

---

year for every 4,000 workers)?"
They are asked to circle one of thirty-seven different values ranging from \$0 to \$6,000 with an open ended "more than \$6,000."

Before turning to the amendments to this process to introduce information from QALY responses, some details of the implementation should be highlighted. First, there are three levels of the probability of accidental death identified. This format is used because the average level associated with job risks need not be the same as the baseline (P_0) and reduced risk (P_1) in a contingent valuation study. Equally important, the *W* and *m* used for the labor supply need not correspond to those associated with the contingent valuation results. They can be specific to each study and to the extent possible should be internally consistent with the information provided by each source.

Finally, if there is one set of estimates for each of the arguments, then the exercise is a calibration or numerical solution that best approximates the points defining the outcomes observed. As more results are available, the process becomes one of estimation. Equations (17a) through (17c) can be interpreted as moment conditions. Several of the variables (e.g. wages and h) are endogenous so that a set of instruments will be needed. This is beyond the scope of this paper, but is certainly an issue to be considered with additions to the base of existing studies providing benefit measures.

3.3 Calibration with QALYs

Our proposal to incorporate QALYs into the preference calibration logic follows the format used with other sources of preference information. At the individual level, we do not require the Hammitt/Broome assumptions. These are largely required to use the QALY index to aggregate health conditions across individuals (or to treat the health condition and its duration for one individual in a specific way). To use QALY responses in preference calibration we simply need to define the QALY question using our specification of preferences, identifying how the health conditions are hypothesized to contribute to the indirect utility function. In this section we consider a Cobb-Douglas direct utility function to illustrate the logic. Following this discussion, in the next section we use the indirect utility function to illustrate how calibration for QALYs might be linked to the calibration for responses to risk. Finally, in the last subsection we consider a further complication. Here we allow health effects to enter preferences in a more complex way, acknowledging that the simple indirect function overlooks the irreplaceability issues and potential relationship between health and the marginal utility of income discussed in the previous section.

3.4 Calibration of QALY Choices without Job Risk

O'Brien and Viramontes [1994] compared QALY and WTP elicitation of health related preferences using an example of chronic lung disease. To motivate full integration of QALY elicited responses with our calibration of job, CV, and labor supply choices involving risks, we consider how their results could be used to calibrate a Cobb-Douglas function. To develop this logic we use their standard gamble question. One option was to remain in current health (H^*) for the remainder of one's life. The other was a gamble between immediate death (H_0) or healthy lung function (H_1) for the rest of life. Their willingness to pay question asked respondents for their maximum WTP for a treatment that had a 1% change of immediate death and a 99% chance of healthy lung function. Suppose preferences could be described with a simple Cobb-Douglas utility function with respect to income (*m*) and health (*H*).

(18) $U(m,H) = m^{\alpha} H^{\beta}$

This form implies that the marginal utility of income is positively related to health and the marginal utility of health is positively related to income. It also implies "standard gamble invariance" with respect to income (i.e. p_s for any health state, H^*, should not vary with income).[12] This response to the SG question implies that individuals select a probability of survival (π_s), such that the following equality holds:

$$(19)\ U(m,H^*) = \pi_s U(m,H_1) + (1-\pi_s)U(m,H_0)$$

Assuming that utility in the death state is equal to zero (the second term drops out), this expression can be rewritten as:

$$(20)\ U(m,H^*) = \pi_0 m^{\alpha} H_1^{\beta}$$

The response to the WTP question implies that individuals select a WTP for the specified treatment (W) such that the following equality holds:

$$(21)\ U(m,H^*) = (0.99)U(m-W,H_1) + (0.01)U(m-W,H_0)$$

Assuming again that utility in the death state is equal to zero implies that:

$$(22)\ U(m,H^*) = (0.99)U(m-W)^{\alpha} H_1^{\beta}$$

Setting equations (20) and (21) equal to one another, it is possible to derive the following expression for the income elasticity parameter as in equation (23).[13]

$$(23)\ \alpha = \frac{ln(\pi_s/(0.99))}{ln(m-W/m)}$$

Using time-tradeoff questions requires that the duration of a health condition be introduced into the preference function. If we maintain the $q = H * T$ formulation the

---

[12] Duration of illness is not explicitly included in this preference function. However, it is implicitly the same for both the SG and WTP cases (i.e. rest of life). If duration were included as an additional multiplicative term in the Cobb-Douglas specification it would not alter the results.

[13] Using sample means and estimates from O'Brien and Viramontes [1994] - π_s=0.83, W=\$1,356, m=\$27,000 – the calibration result for the parameter is $\hat{\alpha} = 3.437$.

H^{β} then would be replaced by $(H * T)^{\beta}$. However, to illustrate an alternative formulation we include duration as a separate argument, as in equation (24).

$$(24)\ U(m, H, T) = m^{\alpha} H^{\beta} T^{\gamma}$$

In this function *T* represents the duration of the health state. In this case, annual WTP to avoid a chronic health state, H^*, can be expressed as:

$$(25)\ U(m, H^*, T) = m^{\alpha} H^{*\beta} T^{\gamma} = (m - W)^{\alpha} H_1^{\beta} T^{\gamma}$$

As can be seen from this expression, *W* (annual WTP) will not vary with respect to the duration of illness (i.e. number of years with illness). A time tradeoff question asks respondents to choose the number of years in perfect health (T_1) that would provide the same level of utility as T^* years in the chronic health state. This equivalence can be expressed as equation (26).

$$(26)\ U(m, H^*, T) = m^{\alpha} H^{*\beta} T^{*\gamma} = m^{\alpha} H_1^{\beta} (t \cdot T^*)^{\gamma}$$

In this expression $t = T / T^*$ is the utility score derived from the time tradeoff approach. Using this preference function implies that *t* will also not vary with respect to the duration of illness. Combining equations (25) and (26) results in equation (27) for the preference parameters.

$$(27)\ \theta = \alpha / \gamma = \frac{ln(t)}{ln(m - W / m)}$$

This equation implies that the ratio of the two parameters can be calibrated, but they cannot be separately identified. Nonetheless, this could be used to derive a benefit function because the WTP function corresponding to this preference structure is also in terms of the ratio of these two parameters. For example, considering a health condition with a time tradeoff score of t^{**}, the corresponding WTP is:[14]

$$(28)\ WTP^{**} = \left(1 - t^{**(1/\theta)}\right)m$$

3.5 Calibration with Job Risk and QALYs

In the previous section, to use QALY responses to standard gamble questions we assumed that individuals' behavior was consistent with expected utility. Thus, it is reasonable to consider the feasibility of treating QALY responses as a separate choice following the same basic preferences as the one displayed in responding to job risks. There are two issues in making this connection. The first is the baseline probability assumption that is maintained in linking the choices and the second is the role of the health index and duration of the health condition in preferences. We addressed this latter issue indirectly in suggesting that *T* could enter, as hypothesized with QALYs, in a multiplicative format or separately, as in equation (24). We return to it below.

In considering the baseline probability issue, we assume these are two different situations because the QALY data available for our illustration relate to a comparison of two hypothetical situations – a standard gamble and WTP comparison in the O'Brien-Viramontes [1994] case, for example. However, it is entirely possible to follow the extension developed in Smith et al. [2002] and consider QALYs using either a scaling (i.e. two lotteries) or translating specification for baseline probabilities. As in our earlier

[14] To illustrate preference calibration using WTP and time tradeoff (TTO) estimates, we use information from a study by Lundberg et al. [1999]. As part of this Swedish study, 366 patients with psoriasis or atopic eczema (ages 17 to 73) were asked both a TTO question for their illness and a maximum WTP question for a cure for their illness. Based on a bidding game approach, the average annual WTP was $1,420. The average annual income of the sample was $16,284, and the average TTO score was 0.9. Applying these values to equation (27) results in the following calibrated value for the parameter ratio: $\hat{\theta} = 1.156$

analysis, these would change some of the calibrating equations, but not in the overall logic. What is especially important about implementing such an extension is the need to consider the informational requirements on respondents' choices and health conditions in order to be able to specify the elements that are hypothesized to describe each choice situation. For now, we will ignore this complexity and consider two extensions to the preference calibration – linking job risk and QALYs and then calibration with a more complex treatment of the health effect in the next section.

We assume the health condition follows the logic in the simplest form of a QALY index (i.e. $q = H \cdot T$ with H the health condition and T the duration), and assume it enters preferences through the intercept of the labor supply equation (i.e. using our earlier notation $\alpha = \alpha_0 \cdot T \cdot H$).[15] In this setting a standard gamble format for deriving the QALY poses a lottery of ideal health (designated as H^* earlier) versus death in comparison to a less than ideal condition $\overline{H}$ for a specified duration T. For our example, we assume that the longevity or duration is maintained constant at T in both the $\overline{H}$ and the ideal health H^* states. π_s is used to designate the stated probability that would make the individual indifferent between the two possibilities.

With this background π_s is defined implicitly by equation (29).

$$(29)\quad \frac{exp\left(\alpha_0 \cdot \overline{H}T + \beta W\right)}{\beta} - \frac{exp(-\mu m)}{\mu} = \pi_s \cdot \left(\frac{exp\left(\alpha_0 \cdot H^{*}T + \beta W\right)}{\beta} - \frac{exp(-\mu m)}{\mu}\right) + (1-\pi_s)\left(-\frac{exp(-\mu m)}{\mu}\right)$$

[15] We have assumed the function $g(H) = \alpha_0 HT$ to simplify the discussion. Clearly the formulation could be made more detailed. With each additional element in the formulation we introduce more prior restrictions.

The specific form of our indirect utility function allows the π_s to be interpreted exclusively in terms of the health index parameters. Equations (30a) through (30d) outline the logic. (30a) simplifies the form and illustrates how the effect of non-wage income drops out.[16]

$$\text{(30a)}\quad \frac{exp\left(\alpha_0 \cdot \overline{H} \cdot T + \beta W\right)}{\beta} = \pi_s \frac{exp\left(\alpha_0 \cdot H^* \cdot T + \beta W\right)}{W}$$

$$\text{(30b)}\quad \pi_s = \frac{exp\left(\alpha_0 \cdot \overline{H} \cdot T + \beta W\right)}{exp\left(\alpha_0 \cdot H^* \cdot T + \beta W\right)}$$

$$\text{(30c)}\quad \pi_s = exp\left(\alpha_0 \left(\overline{H} - H^*\right) \cdot T\right)$$

$$\text{(30d)}\quad \left(H^* - \overline{H}\right) = -\frac{ln\, \pi_s}{\alpha_0 \cdot T} > 0$$

Since π_s will be less than one $ln\pi_s$ will be negative and the reduced value for the health score is positive. To put this relationship in the usual normalized format we could set $\alpha_0 H^* = 1$ and the index becomes $\left(1 - \frac{\overline{H}}{H^*}\right) = -\frac{ln\, \pi_s}{T}$. With information about the longevity conditions (i.e. T) stated in QALY questions and the probabilities stated in the standard gamble, the QALY results can be readily included in preference calibration, parameterizing the intercept in labor supply a specific way.

While this strategy has clear empirical advantages, it overlooks the irreplaceability issues discussed earlier and thus motivates consideration of a more general formulation of the role of health state in individual preferences.

3.6 Generalizing the Role of Health State in Preferences

[16] We should note that this result is not necessarily a condition we suggest is realistic. It is simply an

One of the issues posed by the indirect utility function associated with the semi-log supply arises because the marginal utility of non-wage income is unaffected by health state. Our review of the Cook-Graham logic suggested that this feature (e.g. changes in the marginal utility of income with health state) could play an important role in people's responses to QALY questions. Indeed, the simple QALY example we developed using the Cobb-Douglas example had this feature. Nonetheless, generalizing the function to allow both non-constant marginal utility of income with health and plausible supply responses adds complexity to all the functions used in calibrating preferences. Thus, a decision on how to deal with health state impacts the other aspects of preference calibration. In this section we consider a form that implies a constant elasticity of labor supply, measures non-wage income with an adjustment for health state, and allows the health state to contribute separately to the level of well being as well. Equation (31) provides the proposed general form for the utility in the state alive and (32) for ex ante conception of death.

(31) $$U^A = \alpha \cdot a(q) + exp\left(\frac{(m - b(q))^{\mu+1}}{\mu+1}\right) exp\left(\frac{\beta W^{\eta+1}}{\eta+1}\right)$$

where: $a(q)$, $b(q)$ = functions describing how health quality impacts well being
μ = income elasticity of supply of labor (in terms of non-wage income)
η = elasticity of supply

(32) $$U^D = \frac{m^{\mu+1}}{(\mu+1)}$$

This formulation implies a constant elasticity of labor supply as given in equation (33).

(33) $$H = \beta W^{\eta} (m - b(q))^{\mu}$$

---

outcome due to this formulation of the problem and motivates the analysis in the next subsection.

The next step in the process is to derive for this function each of the equations used in our calibration. We start by adopting a few shorthand expressions to economize on the notation. These abbreviations are given in (34a) through (34c). Equations (35) and (36) provide the results for the VSL and option price with this form.

$$\text{(34a)}\quad A = exp\left(\frac{\beta W^{\eta+1}}{(\eta+1)}\right)$$

$$\text{(34b)}\quad B = exp\left(\frac{m^{\mu+1}}{(\mu+1)}\right)$$

$$\text{(34c)}\quad C = exp\left(\frac{(m-b(q))^{\mu+1}}{1+\mu}\right)$$

$$\text{(35)}\quad VSL = \frac{B-\alpha a(q)-C\cdot A}{(1-P)\cdot C\cdot A\cdot \beta W^{\eta}}$$

$$\text{(36)}\quad (1-P)exp\left(\frac{(m-W\tilde{T}P-b(q))^{\mu+1}}{\mu+1}\right)+P\,exp\left(\frac{(m-W\tilde{T}P)^{\mu+1}}{\mu+1}\right)\left(\frac{1}{A}\right)$$

$$= P_0\frac{B}{A}+\frac{(P_0-P)}{A}\alpha a(q)+(1-P_0)exp\left(\frac{(m-b(q))^{\mu+1}}{\mu+1}\right)$$

As equation (36) indicates the option price is defined implicitly. We cannot solve for a closed form expression, due to the role of health effects (*b(q)*) in the state alive that do not arise in the ex ante state dead. We can define a closed form expression for ex ante WTP. This result follows because we assume the payment is committed only in the situation where the person is alive (e.g. his estate is not liable for ex ante commitments and he knows it). This relationship is given in equation (37).[17]

[17] This is derived starting from the statement of ex ante indifference defined as follows:

(37) $WTP = m - b(q) - \left[\frac{1}{\mu+1} log \left(\frac{1}{A} \frac{(P_0 - P)}{(1-P)} (B + \alpha\, a(q)) + \frac{(1-P_0)}{(1-P)} exp \left(\frac{(m - b(q))^{\mu+1}}{\mu+1} \right) \right) \right]^{\frac{1}{\mu+1}}$

Finally, to complete the equations used in calibration with QALYs we need to specify how the standard gamble response is to be interpreted and derive an expression for the health index. The general form is given in equation (38) with q_i designated as the impaired health state and duration (e.g. recall $q = H \cdot T$) and equation (39) the solution for π_s. To evaluate whether we can recover the parameters of $a(q_i)$ and $b(q_i)$ we need to formulate in more specific terms what is to be assumed about these functions.

$$(38)\ (1-p)\left[\alpha\, a(q) + exp\left(\frac{(m - WTP - b(q))^{\mu+1}}{\mu+1} \right) \cdot A \right] + p\ exp\left(\frac{m^{\mu+1}}{1+\mu} \right) =$$

$$(1-p_0)\left[\alpha\, a(q) + exp\left(\frac{(m - b(q))^{\mu+1}}{\mu+1} \right) \cdot A \right] + p_0\ exp\left(\frac{m^{\mu+1}}{1+\mu} \right)$$

$$(39)\ \pi_s = \frac{\alpha\, a(q_i) + exp\left(\frac{(m - b(q_i))^{\mu+1}}{\mu+1} \right) \cdot A - B}{\alpha\, a(q) + C \cdot A - B}$$

To illustrate the issues that arise in specifying *b(q)* and *a(q)* consider a simple example. We might hypothesize that:

$$(40)\ b(q) = \varphi \cdot (1 - (H \cdot T / L))$$

where L = remaining life expectancy

H and *T* are consistent with our earlier usage and φ is a parameter. This formulation with H=1 for perfect health would imply that there was no reduction in income. A different specification would be required for *a(q)* and potentially different parameters. This type

---

$(1-P)\left[\alpha\, a(q) + exp\left(\frac{(m - WTP - b(q))^{\mu+1}}{\mu+1} \right) \cdot A \right] + P\ exp\left(\frac{m^{\mu+1}}{1+\mu} \right) = (1-P_0)\left[\alpha\, a(q) + exp\left(\frac{(m - b(q))^{\mu+1}}{\mu+1} \right) \cdot A \right] + P_0\ exp\left(\frac{m^{\mu+1}}{1+\mu} \right)$

of relationship could be explored with contingent valuation estimates of WTP (e.g. evaluating alternative forms of $a(q)$), since they would appear in the WTP as implied by our expression in equation (37).

Overall, these expressions indicate that when we interpret individual responses to risk or to choices involving health, the responses to each set of questions will, under a wide array of preference specifications, be dependent on the baseline conditions assumed for the other component of well being. Risk choices depend on baseline health conditions and health choices on job risks. While this finding is certainly plausible it should also be recognized that it adds to the informational requirements imposed on individual studies in order to consistently control the features and background conditions each the model assumes should influence observed responses.

4. Some Illustrative Example

We develop two aspects of the use of QALY information in preference calibration in what follows. First, we consider respondents to the Health and Retirement Study's survey in wave 5 when both willingness to pay and QALY questions about health conditions were asked of the same respondent. After that, we use data from O'Brien and Viramontes [1994] to illustrate preference calibration.

4.1 Can the Same People Make WTP and QALY Tradeoffs?

In wave 5 of the Health and Retirement Study's (HRS) survey of older adults' time/health and money/health tradeoff questions were asked of a subset of the full panel. The HRS is a national panel of birth cohorts 1932-1941 and their spouses. The first wave was conducted in 1992. Interviews have taken place every two years since then, so wave 5 relates to responses in 2000 when the respondents were eight years older. The HRS surveys collect information to consider three sets of economic issues: explaining the antecedents and consequences of retirement; examining the relationships among health, income and wealth over time; and monitoring work and disability interactions. The baseline interviews included 7,600 households (12,652 individuals).

A set of supplementary questions were introduced in wave 5 to consider how these older adults trade time and money for health. While these data offer the potential to estimate a joint framework using both money and time tradeoffs with health, this is beyond our scope here. We use the introduction to the tradeoff questions asked of 910 respondents to evaluate whether these individuals are equally willing to trade off time or money with health. If they are, then the next step in the process would be to use both sets of information jointly in estimating health preferences. If not, then one might argue there are gains from considering each perspective separately to elicit preferences with those respondents most comfortable with each.

Table 1 provides the text of the introductory questions asked before each type of tradeoff question. Table 2 presents the results of a simple contingency analysis, considering whether the same people are equally able to respond to the two types of questions. Clearly there are striking and significant differences, as the chi square test confirms. A much larger proportion of the group is willing to make money than time tradeoffs. Over one-third of the sample receiving the questions are not willing to make either. The more positive insight that emerges from this comparison is that if we consider a composite strategy that allows both time and money tradeoffs to be used, we add 433 respondents (48% of the sample) to the 16% who can answer both. Thus, this limited information suggests that treating the QALY question format as a source of economic choice information enhances what would be available by focusing exclusively on one or the other approach.

4.2 Illustrative Calibrations

O'Brien and Viramontes [1994] in Table 2 of their paper report estimates of the QALYs for a sample of 102 individuals with chronic airways limitation identified from a clinic register. Fifty-three percent reported they had been diagnosed by a physician for having chronic bronchitis, 45% emphysema, and 53% asthma. The standard gamble format asked respondents to compare two hypothetical situations – therapy option Y as a certain outcome where the individual remains in current health status for the rest of his or her life, and therapy option X involving a hypothetical medication that has an uncertain outcome. In therapy option X either the person would have healthy lung functioning or it would result in immediate death. The probabilities of the outcomes were varied to find the risk of death that each respondent felt made therapy X and Y equivalent. In a separate question, these authors also asked a willingness to pay question with the therapy X holding the probability of healthy lung function at 0.99 and death at 0.01. A bidding game format was used to elicit maximum willingness to pay, given that the patients were told their health insurance would not cover the costs.

Table 3 extracts the information necessary for calibration and reports the conversion of responses from Canadian to U.S. currency. The authors do not report the year the data were collected. Given its publication (and the fact that our objective is illustrative) we simply assumed they referred to 1992. We report the standard gamble probability results for the QALY, willingness to pay and income for the full sample, and selected disease states. By assuming 1992, we can match these data readily to the results used to calibrate job risks. Finally, based on the average age of the population and U.S. life expectancy tables, we assume a duration of twenty years (i.e. the anticipated remaining years of life of this age group in the U.S).

The first column reports the background information taken from Smith et al. [2002]. These data combine information from the HRS, Gegax, Gerking and Schulze [1991], and Smith et al. [2003] to allow calibration of the ex ante preferences as described earlier with VSL, labor supply, and contingent valuation estimates of the option price for a risk reduction. Here we report the results for the 50-55 age category because the O'Brien-Viramontes sample is 55. The next three columns report the elicited probability, π_s, monthly willingness to pay, and income levels for the full sample and two sub-samples from the O'Brien-Viramontes study. We used these data sequentially with the information in the first column to calibrate ex ante preferences based on the semi-log labor supply (i.e. with equations (17a) through (17c) and (30d)). The results for the key parameters are given in the lower panel, along with our original calibrated parameters.

Two conclusions follow directly from the comparison. First, the new information does impact economically relevant parameters. The coefficient for non-wage income, μ, changes with the inclusion of the QALY information, suggesting the evaluation of programs that affect the risks of different health states using a benefit function derived in this way would be affected. The dramatic change in the intercept arises from our re-scaling of the QALY index together with the assumed longevity of this sample. The second conclusion from this illustrative exercise is the insensitivity of the parameters to fairly large changes in the elicited probabilities.

It is impossible to draw general conclusions from this illustrative exercise, given the clear issues in matching the multiple sources of data. We do not know the non-wage income, wage rates, or other information for the Canadian sample to compare with the information used in the risk based calibration. Nonetheless, the exercise does demonstrate feasibility. It also indicates that the addition of QALY information impacts parameters we would have hypothesized to be influenced based on a priori theory. Finally, the insensitivity of calibrated parameters to differences in π_s motivates further work with the more complex treatments of health indexes in preferences that were developed at the close of the previous section.

5. Implications

The primary focus of research seeking to use indexes of health conditions, such as QALYs, in environmental benefit estimation has been on assessing the properties of a willingness to pay per QALY. Hammitt [2002b] found such measures are not constant and require stringent conditions be imposed on individual preferences.

This paper has outlined an alternative strategy – treat QALY measurement as another method for eliciting individual preferences over choices involving health and risk and use the responses together with other information about economic choices involving risk to calibrate ex ante preferences.

We have demonstrated this strategy can work. Moreover, we found it is important to consider a range of ways in which individual preferences are impacted by the severity and duration of health conditions. This next step is an essential component of the research required to operationalize our proposed composite preference calibration method for choices involving risks of non-fatal, health outcomes.

References

Brekke, Kjell Arne, 1997, "The Numeraire Matters in Cost-Benefit Analysis," *Journal of Public Economics*, 64 (April): 117-124

Broome, John, 1993, "Qalys," *Journal of Public Economics*, 50 (February): 149-167.

Cook, Philip J. and Daniel A. Graham, 1977, "The Demand for Insurance and Protection: The case of Irreplaceable Commodities," *Quarterly Journal of Economics*, XCI (February): 143-156.

Deaton, Angus and John Muellbauer, 1980, *Economics and Consumer Behavior* (Cambridge: Cambridge University Press).

Gegax, D., S. Gerking, and W. Schulze, 1991, "Perceived Risk and the Marginal Value of Safety," *The Review of Economics and Statistics*, LXXIII (4): 589-96.

Gerking, S., M. de Haan, and W. Schulze, 1988, "The Marginal Value of Job Safety: A Contingent Valuation Study," *Journal of Risk and Uncertainty*, 1 (2): 185-200.

Hammitt, James K. 2002, "QALYs Versus WTP," *Risk Analysis*, 22 (5): 985-1001.

Hammitt, James K. 2002, "How Much is a QALY Worth? Admissible Utility Functions for Health and Wealth," Unpublished Paper, Department of Health Policy and Management, School of Public Health, May.

Hanemann, W. Michael, 1999, "The Economic Theory of WTP And WTA," in I.J. Bateman and K.G. Willis, eds., *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU and Developing Countries* (Oxford: Oxford University Press).

Moore, Michael J. and W. Kip Viscusi, 1988, "The Quantity Adjusted Value of Life," *Economic Inquiry*, 26: 369-388.

Mrozek, Janusz R. and Laura O. Taylor, 2002, "What Determines the Value of Life? A Meta-Analysis," *Journal of Policy Analysis and Management*, 21 (2): 253-270.

O'Brien, Bernie and Jose Luis Viramontes, 1994, "Willingness to Pay: A Valid and Reliable Measure of Health State Preferences?," *Medical Decision Making*, 14 (July/September): 289-297.

Shogren, Jason F. and Thomas D. Crocker, 1991, "Risk, Self-Protection, and Ex Ante Economic Value," *Journal of Environmental Economics and Management*, 20 (January): 1-15.

Shogren, Jason F. and Thomas D. Crocker, 1999, “Risk and Its Consequences,” *Journal of Environmental Economics and Management*, 37 (January): 44-51.

Smith, V. Kerry, Subhrendu K. Pattanayak, and George Van Houtven, 2002, “Using Preference Calibration for VSL Estimation,” Unpublished Paper, Department of Agricultural and Resource Economics, North Carolina State University, December.

Smith, V. Kerry, George Van Houtven, and Subhrendu K. Pattanayak, 2002, “Benefit Transfer via Preference Calibration: “Prudential Algebra” for Policy,” *Land Economics*, 78 (February): 132-152.

Smith, V. K., M. Evans, H. Kim, and D.H. Taylor, Jr., 2003, “Do the ‘Near’ Elderly Value Mortality Risks Differently?” Working Paper, Department of Agricultural and Resource Economics. Raleigh, NC: North Carolina State University.

TABLE 1: QUESTIONS TO INTRODUCE TIME AND MONEY TRADEOFFS WITH HEALTH

HRS Time/Health Tradeoff

Imagine that you will live for exactly ten more years in your current state of health. Now imagine that you can choose to live those ten years in your current state of health, **or** you can choose to give up some years of life and live for a shorter period in perfect health. Would you be prepared to give up any part of those ten years to live in perfect health?

HRS Money/Health Tradeoff

Imagine that you will live for exactly ten more years in your current state of health. Assuming that your current medical expenses and insurance premiums stay the same as they are now, would you be willing to pay any more every month for additional medical treatment if it allowed you to live those ten years in perfect health?

TABLE 2: CHI-SQUARE ANALYSIS OF WILLINGNESS TO MAKE TIME OR MONEY TRADEOFFS WITH HEALTH

Willing to Make a Time Tradeoff	Willing to Make a Money Tradeoff		
	NO	**YES**	**TOTAL**
NO	332	404	736
YES	29	145	174
TOTAL	361	549	910

$\chi^2(1) = 47.57$ (p-value = 0.00)

TABLE 3: ILLUSTRATIVE CALIBRATIONS USING QALYS, JOB RISK, LABOR SUPPLY, AND CV

		Alternative Health Condition O'Brien-Viramontes Current Disease Severity		
	1	Full Sample	Mild	Severe
Background Data (Smith et al. [2002])				
VSL	6,051,270			
p	6.54×10^{-3}			
m	2200			
W	10.27			
Hours	37.96			
Weeks	50			
WTP	909			
p_0	6.60×10^{-4}			
p_1	4.10×10^{-4}			
QALY Data				
π_s		0.83	0.91	0.78
Monthly WTP				
CD$ (92)		113	56	150
US$ (92)		93	46	124
US$ (91)		90	45	120
Income				
CD$ (92)		27,120	22,400	30,000
US$ (92)		22,428	18,525	24,810
US$ (91)		21,773	17,984	24,085
Calibrated Parameters				
α	3.63	89.10	89.10	89.10
β	0.313×10^{-3}	0.314×10^{-3}	0.314×10^{-3}	0.314×10^{-3}
μ	-0.287×10^{-1}	-0.385×10^{-1}	-0.385×10^{-1}	-0.385×10^{-1}
QALY	---	0.99986	0.9999	0.99986

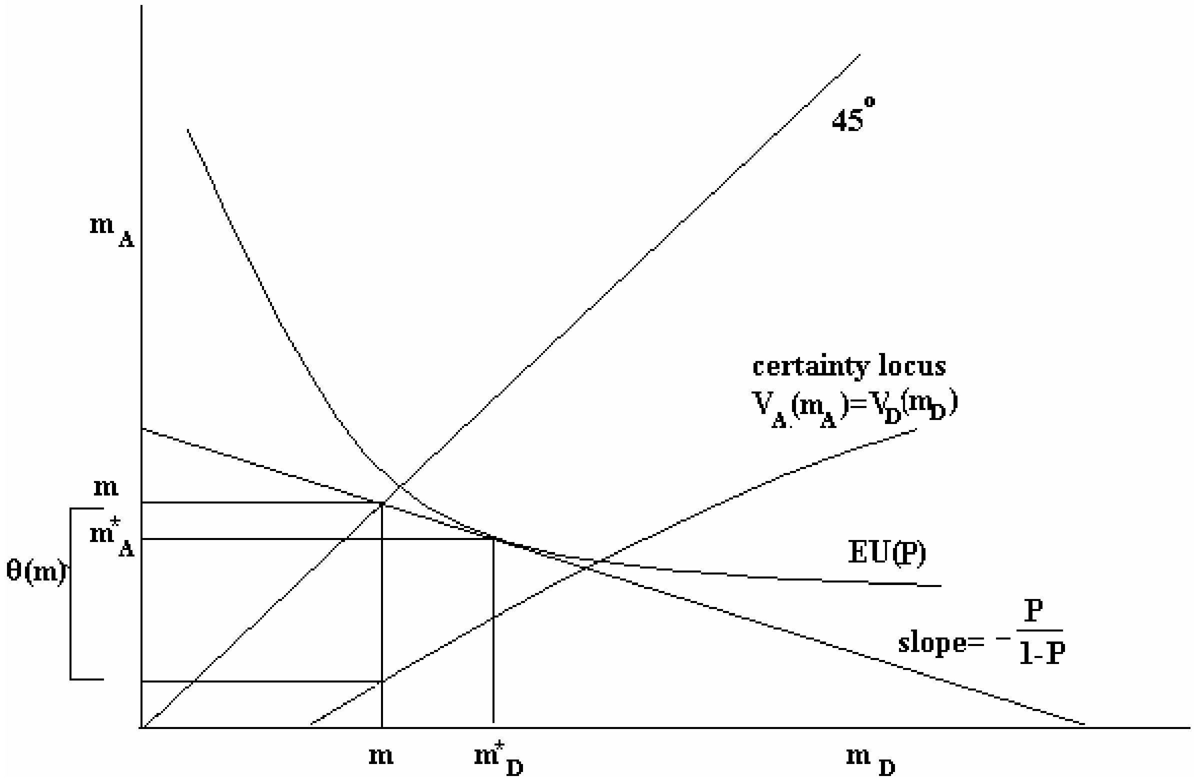

FIGURE 1: EXAMPLE OF EX ANTE TRADEOFFS AND ADJUSTMENT

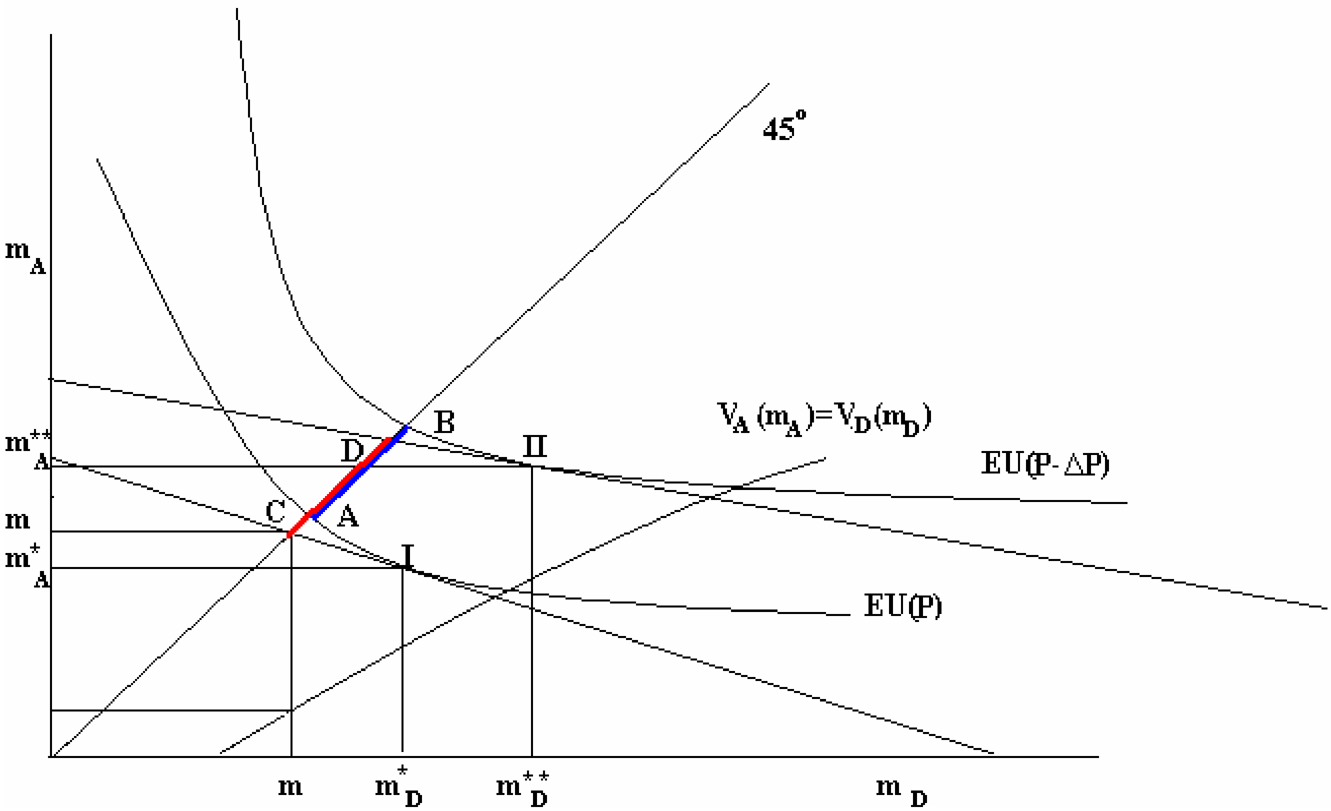

FIGURE 2: ILLUSTRATION OF VALUATION CONCEPTS