Using Preference Calibration for VSL Estimation

V. Kerry Smith
Subhrendu K. Pattanayak
George Van Houtven

Working Paper 03_03 **March 2003**

RTI International's* Environment and Natural Resource Economics Program circulates these working papers to convey information and entertain discussion. The papers may not have undergone formal peer review. Contact Subhrendu Pattanayak (subrendu@rti.org) if you have any questions or concerns. <www.rti.org/enrepaper/>

*RTI International is a trade name of Research Triangle Institute.

Using Preference Calibration for VSL Estimation*

V. Kerry Smith†
Subhrendu K. Pattanayak‡
George Van Houtven§

JEL Classification: Q2
Keywords: Benefit Transfer, Non-market Valuation, Value-of-Statistical-Life

ABSTRACT

To evaluate the benefits of public health and safety programs, considerable attention has focused on the measurement of the economic value of reducing risks to life. The benefits of reductions in mortality risks are typically evaluated using a "unit price" for avoided premature deaths– i.e., the value-of-statistical life (VSL) estimates – which are mostly drawn from wage-risk hedonic studies. In this paper we propose an alternative benefit transfer approach -- preference calibration – for valuing mortality risks. We demonstrate how wage-risk estimates can be supplemented with information from contingent valuation and labor supply studies and linked to a common preference structure. By combining information in this way, it is possible to calibrate parameters of the preference function and develop a WTP function for avoided mortality risks that is inherently more consistent with utility theory.

* This paper was originally presented in January 2003 at the Annual AEA/ASSA Meetings in Washington, D.C.. Financial support for this research was provided by the U.S. Environmental Protection Agency under Contract 68-D-99-024 and Cooperative Agreement CR824861-01-0 and by Health Canada under Contract 4500030756. Thanks are due to Bryan Hubbell, Chris Dockins, and Paul DeCivita for comments on aspects of this research.

† North Carolina State University (kerry_smith@ncsu.edu)

‡ Research Triangle Institute (subrendu@rti.org)

§ Research Triangle Institute (gvh@rti.org)

1. INTRODUCTION

This paper proposes a new approach for developing the benefit estimates used in evaluations of environmental policy. Current practice relies on itemizing the effects of a policy and then monetizing each one with unit values. While this logic may seem to parallel Harberger's [1971] approximation of the consumer surplus associated with price changes, it does not have the same properties.[5] In fact, we will argue that the approach has no direct theoretical basis.

Environmental policy is generally motivated by specific objectives (e.g. to protect human health or to provide swimmable waters in all rivers and lakes, etc.). Thus, it is natural to expect that the design and justification for public actions will have measures of outcomes that are associated with each of these goals. They usually correspond to the effects economic analysts are asked to use in estimating the willingness to pay for the policy. However, this objective does not necessarily require itemizing effects and then relying on unit values to construct monetary benefit estimates.[6] Benefits can be measured from a consistent Hicksian willingness to pay function, provided the existing literature is sufficient to identify and calibrate a function describing how the environmental services of interest (and the related "effects" they have for people) as well as any related market goods influence individual well-being.

We illustrated this logic in a recent paper applied to water quality improvements by combining estimates from recreation demand, hedonic property value, and contingent valuation analyses (see Smith et al. [2002a]). Here, we consider a situation where the approximations used in benefit transfer are potentially more important. They involve the valuation of reductions in mortality risk associated with in air quality improvements. In this case, it has been suggested that there are fewer alternatives to the effects/unit values logic. We demonstrate this conclusion fails to consider the potential for a structural description of the behavior used to estimate how individuals value risk reductions.

---

[5]One way to interpret Harberger's logic for a single price change is to consider it as an approximation for the area under a linearized demand function for a price change from P_0 to P_1. This would be (if $P_1 > P_0$).

$$CS = (P_1 - P_0)x_1 + \tfrac{1}{2}(P_1 - P_0)(x_0 - x_1) = \tfrac{1}{2}(P_1 - P_0)(x_1 + x_0)$$

This can also be expressed in terms of an average of expenditure changes on other goods (with m_o = income).

$$I = m_0 - P_0 x_1 \qquad III = m_0 - P_0 x_0$$
$$II = m_0 - P_1 x_1 \qquad IV = m_0 - P_1 x_0$$
$$A = I - II \qquad B = III - IV$$
$$CS = \tfrac{1}{2}(A + B)$$

At a conceptual level we can envision how price changes with fixed income result in pivoting the budget constraint. The resulting change in expenditures offers one means to reflect the utility change resulting from the price change. This is the point of Harberger's logic. Thus, we can use the changes in remaining expenditures to describe consumer surplus. See Smith and Banzhaf [2002] for further discussion of this logic. Non-market goods do not have a similar link to the budget constraint and thus the pricing of quality changes do not have the same intuitive appeal.

[6]The process of using unit benefit estimates derived in other contexts for a new policy is called benefits transfer. It adapts measures of the economic benefits from a change in some environmental resource so that they can be used to assess the economic value of a similar but separate change in a different resource. Unit values are usually averages derived from one or more past studies that may be normalized first by some measure of the amount of the change that was valued.

Our application is especially relevant for current benefit-cost analyses. Many of the U.S. Environmental Protection Agency's most important regulations involve health effects. EPA's recent reports on the 1990 Clean Air Act Amendments (CAAA) in 1997 and 1999 suggest that the majority of the estimated benefits are due to reductions in premature mortality (EPA, 1997, 1999). Ninety-one percent of annual benefit estimates in EPA's *Prospective Analysis* for improvements due to future CAAA mandates stem from mortality risk reductions. Past improvements also display a similar pattern. EPA's *Retrospective Analysis* found that seventy-five percent of the gains attributed to air pollution reductions from 1970 to 1990 were mortality related.

As a consequence, considerable attention has focused on the measurement of the economic value of reducing risks to life.[7] Conventional practice estimates the economic value of reductions in mortality risk based on the compensation workers are willing to accept to assume increased risks of death on the job. These estimates, labeled the VSL (value of statistical life), are usually interpreted as the sum of the incremental values a set of workers would pay to reduce a common risk they face (e.g. accidental death from job hazards) so that the expected deaths declined by one individual.[8] They are an example of the effects/unit value logic. The effect in this case is the reductions in mortality risk, measured as an expected number of deaths avoided by a group of people who experience reduced ambient air pollution. The unit value is the ex ante marginal rate of substitution, which is treated as if it were a constant. In practice, a diverse set of behaviors have been used to attempt to estimate this ex ante MRS and informal reviews (Unsworth, Neumann, and Browne [1992] and Viscusi [1993]) or statistical meta analyses (Kochi, Hubbell, and Kramer [2002] and Mrozek and Taylor [2002]) have been used to attempt to reconcile the results. This paper demonstrates how one framework can be used to reconcile the estimates.

Section two outlines the logic of preference calibration and adapts it to an expected utility framework. We discuss opportunities for using labor supply estimates to inform risk valuation. The third section considers the difficulties posed by allowing for the baseline risk as well as an individual's age in VSL estimation. The EPA analyses referred to earlier found that the mortality gains occurred primarily to adults over sixty-five years of age. As a consequence, concerns have been raised about the disparity between the ages of the individuals used to measure wage compensation/risk tradeoffs and those affected by policy. Section four discusses the philosophy underlying calibration and the interpretation of the models used for preference calibration as first steps in developing structural models. The last comments on the advantages and disadvantages of the calibration logic.

---

[7]In her letter discussing the Clean Air Compliance Council's evaluation of the *Retrospective Analysis*, Dr. Maureen Cropper, Chair of the Science Advisory Board Committee, highlights the VSL estimates and the disparity between the age distribution of the group used in the measurement versus the one estimated to realize the gains. As a result of this attention, several empirical studies—Krupnick et al. [2002], Alberini et al. [2002], and Smith et al. [2002c]—have considered the age risk valuation issue. It remains a central focus of EPA's environmental research.

[8]See Hammitt [2000] and Freeman [1993] for discussion.

2. PREFERENCE CALIBRATION AND VSL MEASURES

Consistent welfare measures are defined based on relating individual choices to a description of how these decisions arise from budget constrained utility maximization. As a rule, the choices we observe do not offer the ability to characterize fully the underlying preferences. The degree of resolution in each situation will depend both on what has been observed (and recorded for analysis) and on the extent of heterogeneity in preferences that is maintained by analysts modeling people's choices. The logic we propose for benefit estimation relies on analysts being willing to specify all the relevant details that are hypothesized to connect an individual's observed choices to his (or her) underlying preferences and constraints. We will assume the form and the parameters of the underlying preference function are invariant across people. While this restriction is important, it is possible to relax it in several ways and we will illustrate one of them as part of our discussion of how age can influence VSL measures. In our initial applications of the calibration logic for benefit transfer we have assumed that all heterogeneity arises from observable features of the individual. These could include household income or demographic characteristics.

Models describing the wage/risk tradeoff focus on an individual selecting among an array of jobs, each with predefined characteristics that include the working conditions, the risk of fatal accidents, and the compensation. With a continuous array of alternatives, at varying risks, and each individual's decision motivated by attempts to maximize expected utility (as given in equation (1)), the resulting tradeoff between risk and compensation can be described in equation (2).

$$EU = (1-p)U^A(W) + pU^D(W) \tag{1}$$

where

p	=	probability of a fatal accident on the job
W	=	wage rate
$U^A(\cdot)$, $U^D(\cdot)$	=	state dependent utility functions conditional by the outcome alive (A) and dead (D)

$$MRS_{Wp} = \frac{dW}{dp} = \frac{U^A - U^D}{(1-p)U_W^A + pU_W^D}$$
$$U_W^j = \frac{\partial U^j}{\partial W} \tag{2}$$

In this framework job choice is a selection from alternative lotteries. While a clear application of the hedonic logic, this description of the choice process offers few clues as to how to select functional relationships for $U^A(\cdot)$ and $U^D(\cdot)$.

Ordinarily in describing a work/leisure decision it would seem natural to look to specifications that arise in modeling labor supply. However, we were unable to find many applications where

this connection was used to motivate VSL estimates. The only explanation we could locate is in an early discussion of the model by Viscusi [1979]. He discusses labor/leisure choices as part of a general treatment of time allocation and concludes that relaxing the assumption of a fixed hours worked in each of the job alternatives being evaluated did not alter the implied tradeoff between wages and risk. Since this is the primary focus of the model, further attention to these issues with hedonic models has been limited.[9]

Nonetheless, models of labor supply and job choice are not incompatible alternatives. One could, for example, envision a choice setting where workers select among jobs, then conditional on that choice, evaluate the number of hours to be supplied. This perspective has some advantages because it offers another set of information (e.g. labor supply information) for calibrating preferences. As a result, we selected a preference function relying on this basic logic. Following the framework suggested in Burtless and Hausman [1978] we specify a labor supply function and then derive the indirect utility function consistent with it.[10]

A semi-log labor supply, as in equation (3), was selected. The corresponding indirect utility function is given in equation (4).

$$ln(h) = \alpha + \beta W + \mu m \tag{3}$$

$$U^A = \frac{-exp(-\mu m)}{\mu} + \frac{exp(\alpha + \beta W)}{\beta} \tag{4}$$

where

m	=	non-wage income
W	=	wage rate
H	=	hours worked[11]

This specification for U^A is an example that guarantees a simple, observable, labor supply function. We could have started with a more complex indirect (or a direct) utility function. Or, we could add arguments to equation (3). These alternatives simply complicate the algebra but do not, in principle, preclude numerical calibration. With each addition of observable heterogeneity we add associated parameters and expand the demands on the available empirical literature for choice related information. An alternative approach, which we illustrate below, keeps the parameter set small and then calibrates based on different demographic groups using the information available for each group. This strategy implicitly suggests we may not know enough

[9] In an appendix he does suggest that models allowing a marginal time allocation lead to greater attention to the other sources of risk that each individual faces.

[10] Burtless and Hausman [1978] actually used a double log specification. A number of alternatives are possible. Our intention here is to illustrate the general logic so we selected a form that simplifies the algebra. Calibration generally requires a numerical solution of a set of nonlinear equations. Thus, more complex specifications can certainly be considered as long as the parameters to be recovered can be identified.

[11] 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.

to describe the role of specific demographic variables for individual in preferences within a single unified relationship.

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, $U^D(\cdot)$ likely reflects bequest motives. The function would certainly not include the wage rate, but could be related to available non-wage income. In our example we deliberately selected a simplifying assumption, but as with the other decisions associated with our calibration example, this assumption is not required for preference calibration. If $U^D(\cdot)$ corresponds to the first term in the right side of equation (4) (dropping the contribution arising from labor/leisure choices) then the expected utility function is simplified. Equation (5) defines the resulting expected utility and (6) provides the expression for the ex ante MRS, which also corresponds to the VSL.

$$EU = (1-p)\left[\frac{-exp(-\mu m)}{u} + \frac{exp(\alpha + \beta W)}{\beta}\right] + p\left[\frac{-exp(-\mu m)}{\mu}\right] \tag{5}$$

$$-\frac{EU_p}{EU_W} = \frac{dW}{dp} = VSL = \frac{1}{(1-p)\beta} = \frac{W}{(1-p)\eta} \tag{6}$$

where

η = labor supply elasticity

$EU_j = \frac{\partial EU}{\partial_j}$

Equations (3) and (6) illustrate the types of connections between choices and preferences used in preference calibration.[12] The first implies the parameters of labor supply should be related to VSL estimates. Equation (6) describes the specific nature of the connection and implies that estimates of the labor supply elasticity along with an estimate of the wage and the job risk (*p*) are sufficient to infer an implied VSL. Table 1 uses a few estimates from the literature to illustrate this logic. A review of the results in the last column suggests that the range of VSL estimates implied by the labor supply elasticities (together with mean values for the wage rate and job risks) is consistent with what has been found for direct estimates using hedonic models (see Viscusi [1993] and Mrozek and Taylor [2002]).

While this result is broadly supportive of the calibration logic, our purpose is actually to reverse the logic underlying Table 1. That is, we propose to use estimates of the VSL, together with labor supply information and other estimates of how people evaluate risk/money tradeoffs, to recover

[12]These are the types of models used in joint estimation, such as the scheme Cameron [1992] initially proposed.

"estimates" of the preference parameters. For our example, there are three parameters (α, β, μ) and the analysis to this point has considered two choices – job selection and hours worked. A third source is needed for this third parameter when only one estimate of each relationship is available.

To address this issue we use one of the contingent valuation studies of risk/money tradeoffs. We rely on the Gegax, Gerking, and Schulze [1991] study because the risks are described as related to workplace activities. They report willingness to pay (and willingness to accept) estimates for non-marginal changes in risk. To use their results for calibration we need to define this ex ante WTP using our preference function. It is an option price – the non-wage income an individual would be willing to give up to receive the lower risk described in their contingent valuation question.[13] Equation (7) defines this relationship, with p_0 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 $W\tilde{T}P$.

$$\begin{aligned}&(1-p_0)\left[\frac{exp(\alpha+\beta W)}{\beta}-\frac{exp(-\mu m)}{\mu}\right]+p_0\left[\frac{-exp(-\mu m)}{\mu}\right]\\&=(1-p_1)\left[\frac{exp(\alpha+\beta W)}{\beta}-\frac{exp\left(-\mu\left(m-W\tilde{T}P\right)\right)}{\mu}\right]+p_1\left[\frac{-exp\left(-\mu\left(m-W\tilde{T}P\right)\right)}{\mu}\right]\end{aligned} \tag{7}$$

Rearranging terms we can define the option price in equation (8).

$$W\tilde{T}P=\frac{1}{\mu}ln\left[1+\frac{u}{\beta}exp(\alpha+\beta W+\mu m)(p_0-p_1)\right] \tag{8}$$

Equations (3), (6), and (8) define the observed "choice relations." They involve the same types of "results" routinely used in developing unit benefit measures for transfers. There are two important differences in how they are used for preference calibration. First, the results are related to a consistent behavioral function. Thus, the logic does not use a $W\tilde{T}P$ estimate per unit of risk (i.e. $W\tilde{T}P/(p_0-p_1)$). Instead, it defines $W\tilde{T}P$ for a specific expected utility function and assembles information necessary to calibrate the unknown parameters implied by that function. Second, and equally important, the logic recognizes that not all the required information may come from a single study. As a result the interpretation of each choice equation must be made compatible with the other equations (e.g. they should be measured in dollars of the same year) and the

[13]The Gegax, Gerking, and Schulze (1991) 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
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."

components for each relationship should be individually consistent. The W and m relevant to a labor supply response from one study may be different from what is used in the contingent valuation study providing the $W\tilde{T}P$ estimate. Equally important, to the extent we hypothesize that the parameters change nonparametrically with different groups, we could calibrate different sets for different demographic groups. This approach avoids assumptions about how age, race, or gender interacts with the wage, non-wage income, or the job risk in influencing each "type" of individual's choices.

To illustrate the preference calibration logic directly, we select values for the "observables," substitute them into equations (3), (6), and (8), and then solve for the unknown parameters. For example, using data from two related studies (Gegax, Gerking, and Schulze [1991] and Gerking, de Haan, and Schulze [1988]), and the ratios of non-wage to wage income from the National Income and Product Accounts, we can illustrate the calibration process for α, β, and μ. The data used for this calibration include: (a) h = 43.94 hours, (b) W = \$11.76, (c) m = \$1468, (d) VSL = \$1580544, (e) p = 0.00086, (f) WTP = \$655, (g) p_0 = 0.00066, and (h) p_1 = 0.00041. All but m are taken from Gegax, Gerking, and Schulze [1991] and Gerking, de Haan, and Schulze [1988], or supplemented from the original survey, and are in 1983 dollars.[14] The implicit equations to be solved are given in (9a) through (9c).

$$ln(43.94) = \alpha + 10.16\beta + 1468\mu \tag{9a}$$

$$1{,}580{,}544 = (43.94 \cdot 47.64)/(1 - 0.00086)\beta \tag{9b}$$

$$exp(665\mu) = 1 + (\mu/\beta) * 43.94 * (0.00066 - 0.00041) \tag{9c}$$

The resulting calibrated parameters are: α = 3.7694, β = 0.0013, μ = -0.1207. For comparison, with Burtless and Hausman we must convert β and μ into elasticities at the points used for calibration $\eta = \beta \cdot W = 0.015$ and the non-wage income elasticity of labor supply, ϕ, $\phi = \mu \cdot m$ = -177.19. Both are substantially larger in absolute magnitude than their findings (η_{BH} = 0.00003 and ϕ_{BH} = -0.0477). However, these results rely on the subjective risk assessments and the VSL estimates implied by them.

If we considered the matched BLS job risk (p = 0.0834) and corresponding VSL estimates for the same sub-sample (i.e. VSL = 11,837,610 in 1983 dollars) and recalibrate, the parameters would be α = 3.7806, β = 0.000193, μ = -0.0176. The resulting elasticities η^* = 0.0023 and ϕ^* = -25.84 are substantially smaller and in the case of the labor supply, closer to labor supply elasticities reported in Table 1.

---

[14] With the assistance of Shelby Gerking we were able to obtain the original survey data and computed the mean values for hours, weeks worked, and income for respondents who received the willingness to pay version of the questionnaire. The average wage rates were reported in Gegax, Gerking, and Schulze (1991). Non-wage income was approximated by multiplying the annual wage income for the survey respondents answering the WTP question (20, 521) by the sum of rental income and profit relative to total wage income in 1983 (based on the national income and product accounts, 7.15%).

Nonetheless, there remain reasons why the large discrepancy should not be surprising. First, not all of the data used in the calibration are fully compatible. In particular, we do not have a clear basis for estimating the non-wage income. This measure influences both the labor supply and the option price equations. As a result, this comparison should not be considered a refutation of the calibration logic. Rather, it reflects an *opportunity* provided by the preference calibration logic. That is, by requiring benefit estimates be linked to a consistent description of preferences (through the choice relationships implied by those benefit estimates) we have the opportunity to consider other economic interpretations implied by those estimates. In conventional approaches to benefits transfer these comparisons would not be possible.

Indeed, such consistency checks reflect a difficulty with the objective functions raised in some forms of estimation. As we discuss below, in some cases they are too narrow and fail to incorporate all the ways models might be used. The VSL estimate used in our second example was considered to be too large by the original authors. They interpreted the discrepancy as reflecting higher incomes in their sample as compared to the incomes of workers in earlier studies as well as discrepancies between technical and subjective measures for risk. Calibration to a consistent preference structure offers the opportunity to be more systematic in evaluating all the potential implications of a set of estimates.

Before turning to some additional examples, three adaptations to the calibrating equations should be noted. Equation (9b) includes the hours worked in a year in the numerator used to define the VSL. Our structural equation for the VSL (equation (6)) was expressed in terms of the ex ante MRS on an hourly basis. The available VSL estimates are expressed in terms of annual compensation, assuming the individual considers risks and compensation per year. Thus, we scaled the hourly rate on the left side of equation (9b) by the estimated hours worked used in (9a) and an assumption of fifty weeks worked per year to be consistent with the VSL estimate on the left side of the equation. Equation (9c) rearranges equation (8) to simplify the nonlinearity – scaling the Gegax et al. option price estimate by • – and replaces the labor supply equation ($exp(\alpha + \beta W + \mu \cdot m)$) with the level of the labor supplied in a week to be consistent with the timing implied in (9a).

Table 2 illustrates how this logic can be extended to consider what might be described as non-parametric calibration. We combine estimates of the VSL derived for each of three age groups – 51-55, 56-60, and 61-65 – based on the Health and Retirement Study (HRS) in 1991 dollars, reported in Smith et al. [2002c] with the Gerking et al. WTP estimates adjusted to 1991 dollars. In this case we assume that the VSL, baseline probability, non-wage income, wage rate, and hours worked reported from the HRS are relevant for both the labor supply and VSL equations (i.e. equations (3) and (6)). This assumption is especially important for *m* because it is possible to derive a household level estimate from the survey and use it in calibration. For the option price equation we assume the hours worked correspond to the average reported by Gerking et al. (i.e. *h* = 43.94) and assume the $W\tilde{T}P$ would apply to each of these age groups. Now the corresponding parameter estimates imply more plausible labor supply elasticities and still quite large (in absolute

magnitude) non-wage income elasticities. For this group, however, a large negative response may well be more plausible than for the age group considered in Burtless and Hausman. Overall, these computations demonstrate it is possible to use multiple sources of data on labor supply, VSL estimates, and CV measures of the option price individuals would pay to reduce risks of premature death in a behaviorally consistent framework.

Finally, we should note that other behavioral links could have been used. Our analysis began with a model describing labor/leisure choices. So labor supply information became the added component of the information used in calibration. This selection is not essential. We could have considered other forms of mitigating behavior that described an expenditure/risk reduction choice.[15]

3. BASELINE RISK AND AGE

One of the most important discrepancies between policy needs and empirical results involves measuring the benefits relevant for those expected to experience the risk reductions from air quality improvements. Based on EPA's estimates, most of the benefits accrue to people over 65 years of age. Labor market studies of wage/risk tradeoffs relate to middle age workers. This discrepancy has led to a variety of questions and proposed adjustments for benefits transfer. Given the importance of the issue for policies that relate to health risks, it seems natural to ask whether preference calibration can do better than the proposed simple, but ad hoc, adjustments.

At this stage, full calibration to the inter-temporal optimization envisioned in Shepard and Zeckhauser [1984] or Johansson [2001] is beyond our scope and, for practical purposes, infeasible. We do not have the information necessary to fully characterize the choice process and constraints. As a result, in this section we consider the prospects for incorporating age through the assumed baseline risk of death each person perceives when considering the choice of job related risks as further threats to life. Two alternative specifications for the role of baseline risks will be discussed and preference calibrations illustrated in each case. The first involves treating baseline and job related risks as separate lotteries that affect survival probabilities. The second assumes that an individual considers the choice within a reconstituted lottery with a different survival probability reflecting both sources of risk. The first assumes job risks scale the baseline survival probability. The second assumes they translate the risks (Evans and Smith [2002]).

The scaling format arises with Eeckhoudt and Hammitt's [2001] recent reconsideration of issues originally discussed by Sussman [1984]. In their framework, selecting a job with risk p, scales the background risk. That is, each process is a separate lottery. To experience the job risk one must "live" to enter the workplace. Or if the source of background risk is travel, then the process of travel must be survived before the next job related risk is relevant. In this context, expected utility

[15] Quiggin's [2002] recent adaptation of his earlier work on self protection to a state-claims format seems especially well suited for adaptation in these types of applications.

is given in equation (10) with *(1-q)(1-p)* the probability of survival and *(1-(1-q)(1-p))* the probability of death.

$$EU = (1-q)(1-p)\left[\frac{-exp(-\mu m)}{\mu} + \frac{exp(\alpha+\beta W)}{\beta}\right] + (1-(1-q)(1-p))\left[\frac{-exp(-\mu m)}{\mu}\right] \tag{10}$$

The total differential used to define the ex ante MRS (or VSL) is not affected by the level of baseline risk. It contributes to both EU_p and EU_W and thus cancels to yield the original equation (6) for the VSL. The total differential is given in equation (11).

$$\begin{aligned} dEU = (1-q)&\left[\frac{-exp(\alpha+\beta W)}{\beta} + \frac{exp(-\mu m)}{\mu} - \frac{exp(-\mu m)}{\mu}\right]dp \\ &+ (1-p)(1-q)\left(\beta\frac{exp(\alpha+\beta W)}{\beta}\right)dW = 0 \end{aligned} \tag{11}$$

Simplification of equation (11) establishes the basic result for the VSL. It is clear that *(1-q)* can be eliminated without changing *dW/dp*. However, the formulation raises a separate issue, implicit in Sussman's original argument and discussed in detail by Evans and Smith [2002]. The <u>ex</u> <u>ante</u> MRS does vary with age, even in this static formulation if *q* is a function of age. The challenge is to incorporate this effect through the calibration of preference parameters.

Baseline risks do influence the <u>ex</u> <u>ante</u> WTP as seen in equation (12).

$$W\tilde{T}P^* = \frac{1}{\mu} ln\left[1 + \frac{\mu}{\beta} exp(\alpha+\beta W) + \mu m(1-q)(p_0 - p_1)\right] \tag{12}$$

Thus, if we hypothesize that *q* is a function of age, then a three equation system composed of (3), (6), and (12) will allow calibration that accounts for age through the specification of *q* and its role on $W\tilde{T}P$. Panel A in Table 3 re-calibrates preferences using the data in Table 2 with age specific background survival probabilities taken from the Center for Disease Controls' (CDC) national vital statistics. The resulting calibrated parameters are quite comparable to those in Table 2, suggesting that if we believe age affects how people perceive their survival prospects, and adopt the scaling model, then there is little apparent effect on the benefit estimates based on the existing data for different age groups.

The alternative perspective, that Sussman's paper was reacting to, treats the choice process as involving a single reformulated lottery where individuals treat the sources of fatality risk as raising the "total" odds of a fatality.[16] This formulation does impact the VSL but not the <u>ex</u> <u>ante</u> WTP. The revised <u>ex</u> <u>ante</u> MRS is given in equation (13).

[16] It was originally discussed in Freeman [1979] and underlies the approach used in some risk/risk contingent valuation studies. See Magat, Viscusi, and Huber [1996] as one example.

$$\frac{dW}{dp} = VSL = \frac{1}{(1-q-p)\beta} \tag{13}$$

Panel B uses this equation together with (3) and (9) to calibrate preference parameters. The results are given in panel B of Table 3. As in the case of scaling, the calibrated parameter estimates derived using the translating formulation are largely unchanged from the original calibrated values in Table 2.

This relative constancy is potentially important because it suggest that we should not be surprised by recent empirical evidence that finds little difference in risk tradeoffs with age.[17] Two quite different perspectives on the effects of age (through baseline risk) on ex ante marginal rates of substitution cannot be distinguished. That is, translating and scaling have opposite implications for the change in VSL with age – VSL decreases in age with scaling and increases under translating (Evans and Smith [2002]). Yet, our calibration results imply that at the levels of job risk and baseline conditions relevant for the HRS respondents we are unable to discriminate between them.

By contrast, efforts to commoditize the risk/wage tradeoff by redefining the event as the loss of a present value of remaining life years implies a very different adjustment to the VSL and corresponding calibration of the labor supply parameters. Table 4 illustrates these computations for the three age spans in Tables 2 and 3 using the Moore-Viscusi discount rates and life expectancies at the midpoint of each age range. The implied values for β display substantial differences. However, the adjustment in the definition of the event at risk is not consistent with the expected utility logic. It assumes instead that a risk reduction equivalent to reducing expected deaths by one should be measured in terms of the present value of life years for the group experiencing the risk reduction.

4. CALIBRATION AND ESTIMATION

One of the reasons Harberger's approximation for the consumer surplus remains a part of discussions of current methods for policy evaluation is the ability to link his measure to well defined theoretical concepts (see footnote #1).[18] Current approaches to benefit transfer, whether using unit benefits or benefit functions, do not have the same status. While enhancements using statistical summaries no doubt improve our ability to characterize results in the literature, they do not change this basic conclusion. We have proposed a different strategy for benefit transfer. It imposes strong restrictions to assure there is consistency in the use of diverse information. Our example maintains that all individuals have the same preference function and common parameters. This strategy is adopted so the available benefit information can be interpreted within one framework and benefit measures derived from the numerically calibrated function.

[17] Several studies on both CV (Alberini et al. [2002] and Krupnick et al. [2002]) and hedonic wage (Smith et al. [2002c]) have found that VSL estimates do not consistently decline with age as conjectured in policy decisions.

[18] See Hines [1999] for further discussion.

There are at least four advantages to this strategy. First, it assures consistency of the transfers to a unifying economic structure and a well-defined benefit function. Willingness to pay can never exceed income. Second, there are observable "predictions" that can be used to gauge indirectly the plausibility of the benefit function. Our comparison with supply elasticities was one such example. Third, often there are multiple benefit measures available for the resource, risk, or policy outcome of interest. They can arise from models describing different types of choices. Our approach can evaluate their mutual consistency or define conditions required for that consistency. By defining each method's benefit estimate within a single unifying description of an individual choice process (e.g. constrained utility maximization) there is a discipline that forces some reconciliation. Finally, the baseline conditions relevant to choice such as individual income or demographic features (if they are part of the preference model or constraints) must be taken into account.

No doubt this description resembles the background that would be used to characterize an estimation process, except in most cases we often do not have sufficient data to treat each set of estimates as a data point. If we did, then this strategy might be labeled structural meta analysis.[19]

Calibration has been used in several different areas in economics. Initially its use in parameterizing computable general equilibrium models was controversial (see Dawkins, Srinivasan, and Whalley [2001]). In stochastic general equilibrium models such as Kydland and Prescott [1982] reviewers have described their approach as theory with numbers. Dawkins, Srinivasan, and Whalley [2001] describe the primary rationale for calibration in general terms, noting that:

> "The driving forces behind the use of calibration in economics is the belief that any counterfactual analysis is impossible without coherent theoretical framework and that models which are consistent with economic theory are the place to start" (p.3656).

What is at issue in our proposed application of calibration is closer to Hansen and Heckman's [1996] discussion. A statistical summary of available results, expressed as some type of benefit measure, fails to impose a unifying structure that is essential to the ultimate use of that summary. In our case, extrapolation or transfer should be consistent with the process assumed to generate the estimates used in the statistical model. There is nothing in the simple averages or statistical summaries that assures this will be the case.

As Hansen and Heckman [1996] observe, part of the motivation can be overcome by simply developing estimators from different loss functions (p.92). However, there are further requirements – the benefits transferred from existing studies should be consistent with the structural restrictions relevant to these new conditions.[20] It is this consistency condition that

---

[19]This is the logic used in Smith and Pattanayak [2002].

[20]This requirement is different from the use of one set of moment conditions in estimation of a model (what Hansen and Heckman (1996) use to describe step one in calibration) and a second set of moments for testing.

parallels early discussions of the performance of estimators for sets of structural equations based on the predictive performance of the structural economic models versus the fit of each individual equation.

Benefits transfer evaluates policies as counterfactuals. They are only predictions after a decision is made to adopt the policy. Prior to that decision, they are counterfactual analyses of hypothetical outcomes that can never be "checked." As a consequence, consistency requirements are especially important. We will never evaluate the predictions that helped to suggest some policies are misguided. Even if a benefits transfer framework meets acceptable "accuracy" standards in "predicting" measurable gains from a policy once taken, this does not necessarily mean the method was as accurate for the policy that was not selected. A requirement for consistency with the properties implied by a consistent economic model of choice is one approach to assure that assessments of undesirable strategies are more likely to have been credible.

5. IMPLICATIONS

Economic benefits are descriptions of the monetary tradeoffs implied by individual choice. They are not measures people themselves use in making their decisions. The only firm evidence is individual choice. Benefit measures include the details as maintained assumptions. Preference calibration accepts this logic and specifies a specific function to help organize available benefit measures. To the extent the conclusions derived from this process are very sensitive to the specification of preference functions, then it suggests that transfers may need to be limited to situations where analysts can be reasonably confident of the underlying preference function.

In our application of preference calibration to the valuation of risk reductions, there seems to have been a second implication. Policy concerns about the mismatch between the people whose choices yield estimates of the incremental value for risk reductions and the people experiencing the changes may not be warranted. Rather than raising questions about the evidence from both contingent valuation (Krupnick et al. [2002] and Alberini et al. [2001]) and hedonic estimates (Smith et al. [2002c]) suggesting that incremental values do not decline with age, our calibration results indicate at the levels of risk, labor supply, wages, and non-wage income we should not have expected to find differences. Thus, this conclusion casts doubt on simple adjustments relying on the use of value per discounted life year remaining.

The consistency requirements of preference calibration for models where age is treated as reducing the baseline survival probability confirm the CV and hedonic results. That is, large differences in the incremental values were judged as unlikely. Thus, perhaps we should "turn the tables" and question the ad hoc adjustment of VSL per discounted life year as a credible basis for benefits transfers.

The next step in this research is to move beyond summaries of a few point estimates. Our calibration confirms it is possible to use this logic to develop estimates of the preference function from these secondary data—each treated as a sample of "aggregate" statistics. The logic follows Imbens and Lancaster's [1994] call for multiple sample GMM estimation, combining micro and aggregate sources of data.

REFERENCES

Alberini, Anna, Alan Krupnick, Maureen Cropper, Nathalie Simon, and Joseph Cook. 2001. "The Willingness to Pay for Mortality Risk Reductions: A Comparison of the United States and Canada." Working Paper, October.

Burtless, Gary, and Jerry A. Hausman. 1978. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." *Journal of Political Economy* 86(December):1103-1130.

Cameron, Trudy A. 1992. "Combining Contingent Valuation and travel Cost Data for the Valuation of Nonmarket Goods." *Land Economics* 68(3):302-317.

Dawkins, C., T.N. Srinivasan, and J. Whalley. 2001. "Calibration." Chapter 58 in J. Heckman and E. Leamer (eds.) *Handbook of Econometrics*, Volume 5 (Elsevier Science B.V.).

Eeckhoudt, L., and J. Hammitt. 2001. "Background Risks and the Value of Statistical Life." *Journal of Risk and Uncertainty* 23(3):261-279.

Evans, M.F., and V.K. Smith. 2002. "VSL Properties with Multiple Risks: Translating versus Scaling." Working Paper - CeNREP. Raleigh, NC: North Carolina State University.

Freeman, A. Myrick III. 1979. *The Benefits of Environmental Improvement: Theory and Practice* (Baltimore: Johns Hopkins University).

Freeman, A. Myrick III. 1993. *The Measurement of Environmental and Resource Values: Theory and Methods* (Washington, DC: Resources for the Future).

Fuchs, V.R., A.B. Krueger, and J.M. Poterba. 1998. "Economists' Views about Parameters, Values, and Policies: Survey Results in Labor and Public Economics." *Journal of Economic Literature* 36(September):1387-1425.

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. 2000. "Valuing Mortality Risk: Theory and Practice." *Environmental Science and Technology* 34(8):1396-1400.

Hansen, L., and J. Heckman. 1996. "The Empirical Foundations of Calibration." *Journal of Economic Perspectives* 10(1):87-104.

Harberger, Arnold C. 1971. "Three Postulates for Applied Welfare Economics: An Interpretive Essay." *Journal of Economic Literature* 9(September):785-797.

Hines, James R., Jr. 1999. "Three Sides of Harberger Triangles." *Journal of Economic Perspectives* 13(Spring):167-188.

Imbens, G.W., and T. Lancaster. 1994. "Combining Micro and Macro Data in Microeconometric Models." *Review of Economic Statistics* 61(October):655-680.

Johansson, Per-Olov. 2001. "Is There a Meaningful Definition of the Value of a Statistical Life?" *Journal of Health Economics* 20:131-139.

Kochi, I., B. Hubbell, and R. Kramer. 2002. "An Empirical Bayes Approach to Combining and Comparing Estimates of the Value of a Statistical Life." Working Paper. Durham, North Carolina: Nicholas School of the Environment, Duke University.

Krupnick, A., A. Alberini, M. Cropper, N. Simon, B. O'Brien, R. Goeree, and M. Heintzelman. 2002. "Age, Health, and Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents." *Journal of Risk and Uncertainty* 24(2):161-186.

Kydland, Finn E., and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica* 50(November):1345-1370.

Magat, Wesley A., W. Kip Viscusi, and Joel Huber. 1996. "A Reference Lottery Metric for Valuing Health." *Management Science* 42(August):1118-1130.

Meng, Ronald. 1989. "Compensating Difference ins the Candadian Laour Market." *Canadian Journal of Economics* 22(2):413-24.

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

Mrozek, J., and L.O. Taylor. 2002. "What determines the Value of Life? A Meta Analysis." *Journal of Policy Analysis and Management* 21(2):253-70.

Pencavel, John. 1986. "Labor Supply of Men: A Survey" in O. Ashenfelter and R. Layard, eds., *Handbook of Labor Economics, Volume I* (Elsevier Science Publishers BV): 3-102.

Quiggin, John. 2002. "Risk and Self-Protection: A State-Contingent View." *Journal of Risk and Uncertainty* 25(September):133-146.

Shepard, Donald S. and Richard J. Zeckhauser. 1984. "Survival versus Consumption." *Management Science* 30:423-439.

Smith, V.K. and H. Spencer Banzhaf. 2002. "A Diagrammatic Exposition of Weak Complementarity and the Willig Condition," revised working paper, CEnREP, North Carolina State University, October.

Smith, V.K., and S.K. Pattanayak. 2002. "Is Meta-Analysis the Noah's Ark for Non Market Valuation?" *Environmental and Resource Economics* 22(1-2):271-296.

Smith, V.K., G.L. Van Houtven and S.K. Pattanayak. 2002a. "Benefit Transfer via Preference Calibration: Prudential Algebra for Policy." *Land Economics* 78(1):132-152.

Smith, V. K., S. K. Pattanayak and G. Van Houtven. 2002b. "VSL Reconsidered: What do Labor Supply Estimates tell us about Risk Preferences?" Discussion Paper 02_08. Research Triangle Park, NC: RTI.

Smith, V. K., M. Evans, H. Kim, and D.H. Taylor, Jr. 2002c. "Do the 'Near' Elderly Value Mortality Risks Differently?" Working Paper, Department of Agricultural and Resource Economics. Raleigh, NC: North Carolina State University.

Sussman, Frances G. 1984. "A Note on the Willingness to Pay Approach to the Valuation of Longevity." *Journal of Environmental Economics and Management* 11(March):84-89.

Thaler, R. and S. Rosen. 1976. "The Value of Saving a Life: Evidence from the Labor Market," *Household Production and Consumption* N.E. Terleckji. New York, NBER.

Unsworth, Robert, James Neumann, and W. Eric Browne. 1992. Memorandum to James DeMocker-subject "Review of Existing Value of Life Estimates: Draft Valuation Document." Industrial Economics, Inc., November 6.

U.S. Environmental Protection Agency. 1997. *The Benefits and Costs of the Clean Air Act 1970 to 1990*. Report to Congress, October.

U.S. Environmental Protection Agency. 1999. *The Benefits and Costs of the Clean Air Act 1990 to 2010*. Report to Congress, November.

Viscusi, W. Kip. 1979. *Employment Hazards: An Investigation of Market Performance* (Cambridge: Harvard University Press).

Viscusi, W. K. 1980. "Unioin, Labor Market Structure, and the Welfare Implications of the Quality of Work." *Journal of Labor Research* 1(1): 175-192.

Viscusi, W. K. 1993. "The Value of Risks to Life and Health." *Journal of Economic Literature* 31(5):1912-1946.

Table 1. VSL Estimates Implied by Labor Supply Elasticity, Hourly Wage Rate, and Average Job Risk*

Labor Supply Elasticity(η)	Wage Rate (r) (1998 $ / hour)	Job Risk (p) (1 x 10-4)	VSL (1998 dollars)
0.02[a]	18.5[f]	10.98[f]	1,852,034
0.02[a]	19.59[g]	6.87[g]	1,960,347
0.02[a]	15.12[h]	1.18[h]	1,512,178
0.02[a]	27.67[i]	1.9[i]	2,767,526
0.06[b]	18.5[f]	10.98[f]	617,345
0.06[b]	19.59[g]	6.87[g]	653,449
0.06[b]	15.12[h]	1.18[h]	504,059
0.06[b]	27.67[i]	1.9[i]	922,509
0.1[c]	18.5[f]	10.98[f]	370,407
0.1[c]	19.59[g]	6.87[g]	392,069
0.1[c]	15.12[h]	1.18[h]	302,436
0.1[c]	27.67[i]	1.9[i]	553,505
0.14[d]	18.5[f]	10.98[f]	264,576
0.14[d]	19.59[g]	6.87[g]	280,050
0.14[d]	15.12[h]	1.18[h]	216,025
0.14[d]	27.67[i]	1.9[i]	395,361
0.21[e]	18.5[f]	10.98[f]	176,384
0.21[e]	19.59[g]	6.87[g]	186,700
0.21[e]	15.12[h]	1.18[h]	144,017
0.21[e]	27.67[i]	1.9[i]	263,574

*This table is taken from Smith et al. [2002b].

[a]Lowest Marshallian labor supply elasticity from experimental data reported in Pencavel (1986).

[b]Lowest Marshallian labor supply elasticity from experimental data reported in Pencavel (1986).

[c]Consensus Marshallian labor supply elasticity from non-experimental data reported in Fuchs, Krueger, and Poterba (1998).

[d]Highest Marshallian labor supply elasticity from non-experimental data reported in Pencavel (1986).

[e]Highest Marshallian labor supply elasticity from experimental data reported in Pencavel (1986).

[f]Average job risk and hourly wage rate in Thaler and Rosen (1976) as cited in Mrozek and Taylor (2002).

[g]Average job risk and hourly wage rate in Gegax, Gerking, and Schulze (1991) as cited in Mrozek and Taylor (2002).

[h]Average job risk and hourly wage rate in Viscusi (1980) as cited in Mrozek and Taylor (2002).

[i]Average job risk and hourly wage rate in Meng (1989) as cited in Mrozek and Taylor (2002).

Table 2. Combining Contingent Valuation and Age Specific VSL Estimates

	Ages		
	51-55	**56-60**	**61-65**
VSL	6,051,270	6,421,845	10,038,357
p	$6.54.10^{-5}$	$5.81x10^{-5}$	$5.85x10^{-5}$
m	2200	2563	2830
W	10.27	10.09	10.24
Hours	37.96	37.31	35.72
Weeks	50	50	50
WTP	909	909	909
p_0	$6.60x10^{-4}$	$6.60x10^{-4}$	$6.60x10^{-4}$
p_1	$4.10x10^{-4}$	$4.10x10^{-4}$	$4.10x10^{-4}$
Calibrated Parameters			
α	3.633	3.616	3.573
β	$0.313x10^{-3}$	$0.289x10^{-3}$	$0.175x10^{-3}$
μ	$-0.287x10^{-1}$	$-0.267x10^{-1}$	$-0.164x10^{-1}$
η	0.0032	0.003	0.0018
ϕ	-61.6	-66.6	-45.28

Table 3. Incorporating Age Specific Baseline Risks into Preference Calibration: Scaling versus Translating

	Ages		
	51-55	**56-60**	**61-65**
q	5.87×10^{-3}	1.33×10^{-2}	3.25×10^{-2}
VSL	6,051,270	6,421,845	10,038,357
p	6.54×10^{-5}	5.81×10^{-5}	5.85×10^{-5}
m	2,200	2,563	2,830
W	10.27	10.09	10.24
Hours	37.96	37.31	35.72
Weeks	50	50	50
WTP	909	909	909
p_0	6.6×10^{-4}	6.6×10^{-4}	6.6×10^{-4}
p_1	4.1×10^{-4}	4.1×10^{-4}	4.1×10^{-4}
Scaling			
α	3.633	3.616	3.573
β	0.313×10^{-3}	0.289×10^{-3}	0.175×10^{-3}
μ	-0.287×10^{-1}	-0.267×10^{-1}	-0.164×10^{-1}
η	0.0032	0.0030	0.0018
Translating			
α	3.633	3.616	3.573
β	0.315×10^{-3}	0.293×10^{-3}	0.181×10^{-3}
μ	-0.287×10^{-1}	-0.267×10^{-1}	-0.164×10^{-1}
η	0.0032	0.0030	0.0018

Table 4. Quantity Adjusted Life Years and Calibrated Preference Parameter

Age Category	β	Years of Life Expectancy (T*-Age)[a]	VSL	VSL/δ[b]
51-55	1.595×10^{-6}	27.2	6,051,270	626,971
56-60	1.446×10^{-6}	23.1	6,421,845	691,815
61-65	8.735×10^{-7}	19.2	10,038,357	1,144,937

[a]Taken from Table 1 in *National Vital Statistics Report,* Vol. 50, No. 6, March 21, 2002.

[b] $\delta = \frac{1}{\rho}\left(1 - exp\left(-\rho\left(T^* - Age\right)\right)\right)$ with ρ = discount rate. For our calculations we assumed the discount rate was 0.096 (Moore and Viscusi's [1988] estimate), and age at the midpoint of the range.