Economic Valuation of Mortality Risk Reduction: Assessing the State of the Art for Policy Applications

PROCEEDINGS
Session VI
Research Frontiers

A workshop sponsored by the
US Environmental Protection Agency's
National Center for Environmental Economics
and National Center for Environmental Research

November 7, 2001
The Holiday Inn Silver Spring Hotel
8777 Georgia Avenue
Silver Spring, MD 20910

Edited by Abt Associates Inc.
4800 Montgomery Lane, Suite 600
Bethesda, MD 20814

Acknowledgements

Sections of this report indicated as Asummarizations@ were prepared by Abt Associates Inc. with funding from the National Center for Environmental Economics. Abt Associates wishes to thank Nicole Owens and Kelly Maguire of EPA's National Center for Environmental Economics.

Disclaimer

Although the information in this document has been funded in part by the United States Environmental Protection Agency under Contract # 68-W6-0055 to Stratus Consulting Inc., it may not necessarily reflect the views of the Agency and no official endorsement should be inferred.

Proceedings for Session VI
--Table of Contents—

UNDERSTANDING INDIVIDUAL PREFERENCE FOR REDUCTIONS IN MORBIDITY-MORTALITY EVENTS

Presented by J.R. DeShazo, UCLA
Co-authored with Trudy Ann Cameron, UCLA

Summarization

Dr. DeShazo described what he was going to present as a proposal, a project that will soon be funded by Health Canada and USEPA. Noting that it addresses some of the concerns that had already been raised at the workshop, he told the audience that an alternative title that they considered for their research was AFates Worse Than Death.@

Their interest in this project, he said, was stimulated by the tremendous variability that=s in the literature on VSLs. They wanted to try and identify the systematic determinants of the variability in VSLs that is observed across the 26 studies that make up the current VSL estimate or, more broadly, the 60-some studies that exist in the literature. Rather than doing a meta-analysis, they propose to do a very detailed level conjoint state of preference choice experiment.

They start by considering an individual who identifies certain types of morbidity and mortality, certain causes of death that he=d like to avoid. He may choose an intervention to affect the probability that the mortality or morbidity event occurs or the severity with which it occurs. Typically, Dr. DeShazo said, the interventions that we choose are very specific, and individuals choose across interventions. Their research was motivated by the realization that much of the existing literature seems to under-characterize both the outcomes that people care about and the attributes of the interventions that are supposed to bring about a change in the outcome.

The key feature of their study is to specify a fuller utility function and allow individuals= values of statistical life, as expressed in their private behavior, to vary systematically across (1) classes of risk-reducing interventions, (2) the types of premature deaths (i.e., ways of dying) those interventions target, (3) the characteristics of the affected individuals, (4) complements to and substitutes for risk-reductions, and (5) characteristics of the individual=s local risk levels and amenities. Using two conjoint choice exercises in which respondents are asked to choose among three different risk-reduction interventions that mitigate a specific way of dying, they will then estimate the representative utility function.

Dr. DeShazo said that researchers don=t adequately characterize the process of dying, and because of that, they=re probably missing a lot of what people care the most about when they are evaluating changes in morbidity or mortality outcomes. We see routinely in the context of living wills and other kinds of behavior that there are fates worse than death, he said, and people will take lots of precautions to avoid them.

Their study will be a stated preference conjoint study, and they will apply it to a representative random sample of households in the U.S. and Canada. They're hoping to have 4,400 households in their study and to match their survey data with Census tract-level data (or something approximating that) to evaluate geographical characteristics and variability and individual local risk levels for all the prices and amenities that might react with people's willingness to pay.

Conjoint choice experiments are choice settings in which the researchers introduce alternatives, including an alternative "none of the above" so that people can choose to opt out of the choice scenario. Drs. DeShazo and Cameron propose initially characterizing a type of premature death – e.g., cancer, stroke, or heart attack – and randomly varying the attributes associated with the process of dying that way. Then they would introduce the risk associated with that type of death, and finally they would introduce an intervention that would alter that baseline outcome, describing a variety of attributes of that intervention, including the price.

Dr. DeShazo said he would next go through the sets of attributes (given above), noting that they would start with 30 to 35 different attributes (across all five sets), but that they would probably only use 10 to 14 of those in the choice experiment.

On the risk-reducing interventions: Dr. DeShazo emphasized that the scope of the interventions presented in their conjoint exercise is going to vary. Some interventions will have either purely private impacts or scopes or perhaps a private family scope. Other interventions will affect the communal distribution of risk and the communal distribution of morbidity and mortality outcomes. They will try to identify the different preference structures and how they vary over these interventions, as a function of their privateness or publicness. In both cases (private and public interventions), people would be making private financial contributions.

He noted that although interventions usually are characterized solely in terms of one attribute, their price, they really vary along a wide range of characteristics. In addition to the price, that gets modified by the type of payment vehicle, the intervention will vary in terms of how often you have to use it and how long you have to use it. It's also going to vary, importantly, in some non-price cost effects, such as time cost and discomfort. So, for example, taking a hypertension pill is quite different in terms of disutility and time loss than getting an annual mammogram or a proctological exam.

Similarly, many of these interventions have non-risk benefits – that is, they reduce morbidity in a variety of ways, and, again, he noted, this is typically omitted when we're specifying and estimating utility functions. Also, a lot of the interventions are going to vary in their perceived effectiveness. Dr. DeShazo said they want to evaluate people's perceptions of the interventions, both in terms of their scope and their effectiveness and some other attributes.

On the outcome to be affected: Dr. DeShazo said they were still discussing whether or not to explicitly identify a cause of death in their conjoint exercise, although they clearly want to start their discussion with respondents with a focus on the ultimate outcome that they care about.

If they do use types of death, he said, they would use the most common. He noted that about 94 percent of people die of cancer, heart disease, respiratory disease, or stroke. But researchers leave out of their analyses the attributes of the dying process -- the degree and duration of pain, the degree and duration of self-impairment, the extended duration of mental deterioration, and the financial impact on survivors. Dr. DeShazo suggested that when people are undertaking their private investments across various interventions, much of what they're thinking about, in addition to changing the probability that the death is going to happen, is what the absolute levels of these attributes of the death (e.g., the degree and duration of pain) are going to be for them. Interim morbidity reductions that can be achieved through the intervention are also routinely left out of analyses, he said, and can be characterized in terms of avoided illness, avoided hospitalization, avoided injury, and so on. The characteristics that are typically focused on for an outcome, he said, are the baseline risk, the change in the level of likelihood, and any latency characteristics of the outcome **B** that is, how far from the intervention (temporally), do these outcomes express themselves?

The contemporary stated preference (SP) approach doesn't allow us to identify an individual's marginal rate of substitution across any of the attributes of the outcome at the moment, Dr. DeShazo noted, because we omit so many of them. Nor are we able to characterize individuals' marginal rates of substitution across types of interventions. He said that this can be done in the context of a conjoint solicitation method and that they hope that those two characterizations will be two very likely outcomes of their project.

On individual heterogeneity: Dr. DeShazo noted that individual heterogeneity can come from a variety of sources and that Knowledge Networks, their contractor for this project, fortunately already has a lot of information on the subjects' personal characteristics. Because of this, a lot of the survey time can be spent on communicating changes in risk and communicating changes in morbidity.

On complements to and substitutes for risk-reducing interventions: Dr. DeShazo said they also will try and control for complements to and substitutes for the morbidity and mortality risk reduction, using factors such as the individual's mental and health status. He noted that there are many psychometric measures that doctors and health economists have been using to evaluate the discomfort associated with different medical treatments, in terms of choosing alternative therapies. They will try to use some of those in their study.

It is known, he said, that the propensity to retire depends a lot on the health and presence of a surviving spouse and other family members. It would be expected that these factors would also be complements to longevity and quality of life. Health care coverage and life insurance will also affect individuals' values of an intervention.

Dr. DeShazo said that they haven't fully thought out opportunities for utilizing the geographical differences across the U.S. and Canada in the local baseline levels of risk from various sources. He noted, however, that these vary tremendously, and that they're hoping to make use of that variability when considering the hypothesis that these baseline risks affect the

type of intervention and the types of morbidity and mortality that people choose to reduce most. He noted also that the relative prices are going to change and the level of other potential amenities which might be complements will also change geographically.

Finally, he said they are going to have a module in which the individual is asked to undertake a variety of personal assessments to enable them to measure people=s subjective perceptions on various things, including longevity. For some causes of death, Dr. DeShazo noted, individuals don=t have that many interventions to choose from B the controllability of their risk, from the individual=s perspective, is very limited -- while for other forms of death there are lots of interventions and we can substitute relatively cheaply across them, for example through diet, through medical treatment, through exercise, through changing risky behaviors, etc. This module, he said, is meant to get at the subjective assessments of substitutability across interventions, and the Acontrollability@ of interventions in terms of their availability. He noted that researchers often talk about controllability or the voluntary nature of risk exposure. With this module, Dr. DeShazo said, they will be trying to get at people=s perceptions of these attributes. He focused here on dread, in particular, which he referred to as the other poorly defined, very aggregate concept that we often hear talked about in the context of cancer (going back to the characteristics of the dying process). Assuming that dread can be disaggregated in terms of factors such as duration and intensity of pain, and mental impairment, they hope this module will shed light on how dread affects how people value risk reductions.

Extending the Reference Lottery Approach for Valuing Fatal Cancer Risks
A Brief Summary of On-going Research
--Working Paper*--

PRESENTED BY:
GEORGE VAN HOUTVEN
RESEARCH TRIANGLE INSTITUTE

CO-AUTHORS:
CHRIS DOCKINS
US EPA
MELONIE WILLIAMS
US EPA

*This is a working paper developed for the US Environmental Protection Agency National Center for Environmental Economics and National Center for Environmental Research's workshop, "Economic Valuation of Mortality Risk Reduction: Assessing the State of the Art for Policy Applications," held November 6-7, 2001 at the Holiday Inn Silver Spring Hotel in Silver Spring, MD.

Extending the Reference Lottery Approach for Valuing Fatal Cancer Risks
A Brief Summary of On-going Research

Presented at
EPA Workshop: Economic Valuation of Mortality Risk Reduction

Presented by
George Van Houtven (RTI)

with
Chris Dockins (EPA)
Melonie Williams (EPA)

November 7, 2001

Overview

A fundamental problem for applied benefit cost analysis for environmental policy is that there is a mismatch between fatal risks addressed by most valuation studies and fatal environmental risks. Most empirical VSL estimates are derived for immediate, accidental (e.g., job-related) risks where this no delay between exposure and mortality. Consequently, there is no extended period of morbidity associated with measurable mortality risks from the workplace. For many environmental risks, on the other hand, there is a lag between initial exposure and an increase in the probability of dying. Environmental cancers, a major focus of environmental policy, are associated with long latency periods between exposure and onset of illness. Further, cancer is associated with painful and often lengthy periods of morbidity between onset and death. Complete environmental benefits analysis requires consideration of the unique aspects of environmental cancer risks.

To date there have been few attempts to focus original valuation research on cancer risks, but there are several possible approaches available. The first alternative to observe cancer risk - dollar tradeoffs individuals make in markets or constructed markets. There are few natural experiments available for estimation of revealed preferences for reduced cancer risks. While individuals and households undoubtedly make tradeoffs to reduce these risks, it's not always evident what the objective or subjective risk reductions are, nor how individuals consider the timing of exposure and onset in their decisions. Stated preference techniques are complicated by serious risk communication and risk comprehension hurdles (Hammitt XXXX.)

The alternative we consider here is to elicit risk vs. risk tradeoffs from individuals. Because this technique does not necessarily require explicit valuation and comprehension of very small risks, it lowers the cognitive burden for respondents. The analytic tradeoff is that risk-risk

comparisons reveal only a relative value for cancer risk reductions, but this may be exactly what is needed for policy analysis: an adjustment factor to facilitate benefit transfer from existing VSL estimates.

The Reference Lottery Metric

The reference lottery approach for assessing risk-risk tradeoffs begins with asking survey respondents to choose between two hypothetical locations, both of which imply different lotteries regarding three possible health outcomes:

H_0 : current health, with probability p_0
H_1 : bad health state 1, with probability p_1
H_2 : bad health state 2, with probability $p_2 = 1 - p_0 - p_1$

The researcher then can use these choices to infer individuals' relative preferences (marginal rate of substitution) for reducing the risks of the alternative bad health states:

$$MRS_{12} = -dp_1/dp_2$$

A "chained approach" is used to apply the reference lottery results to assess risk values. The value of reducing p_1 (V_1) can be inferred from the value of reducing p_2 (V_2) by

$$V_1 = (1 / MRS_{12}) * V_2.$$

If V_2 is the value of reducing an immediately fatal risk then one can use existing VSL estimates from hedonic wage or transportation studies.

The literature contains a few prior examples of the reference lottery approach for assessing risk-risk tradeoffs. The technique is essentially a simplified conjoint format developed originally by Viscusi et al. (1991) who constructed a tradeoff between automobile accident death and chronic bronchitis. Krupnick and Cropper (1992) also employed the technique to examine relative preferences for reducing chronic bronchitis risks with a focus on the impact of knowledge of the health effect. Magat et al. (1996) first used risk-risk tradeoffs to estimate the relative value of reducing cancer risks. Respondents were also asked about relative preferences for reduced risk of nerve disease. A key omission in Magat et al. is that time is never addressed explicitly in the survey, so there is no way to separate the effect of latency on value.

We extend the reference lottery metric for assessing fatal cancer risk valuation by focusing specifically on how mortality risk values vary with respect to the length of the latency period and the length of the morbidity period. This allows to test directly the theoretical prediction of the life-cycle consumption model that future VSL is equal to the discounted value of current VSL.

Our survey elicits responses for tradeoffs between fatal stomach cancer risks vs. auto death risks. Stomach cancer was chosen for several reasons. First, it is an easily understood type of cancer, with symptoms similar to those of many internal cancers that may be linked to environmental causes. Second, stomach cancer occurs with an annual incidence (baseline risk) on the order of

10^{-4}, which is similar to that of automobile accident fatality. This facilitates comparison of the two types of risks. Finally, stomach cancer has a relatively low survival rate which makes the restriction of attention to fatal risks more plausible.

Mode of administration and Survey Outline

We will use WebTV technology to conduct the survey with an existing panel of users maintained by Knowledge Networks. This is a broad and reasonably representative national sample, making it easier to obtain results applicable for policy analysis. Web-based surveys also offer design advantages. For example, the survey can be interactive and skip patterns are much more easily implemented than with mail surveys.

Survey Outline

$ *Introduction*. Respondents are introduced to the subject area and purpose of the study, including a reminder that people face risk tradeoffs every day. For example, people must often decide:

- whether to have surgery, when there is a chance that serious complications will occur,
- whether to take medications, when there is a chance that they will cause harmful side effects, or
- whether to buy a car with airbags, when there is a chance that airbags can cause certain injuries.

$ *Health status.* Several questions are directed at eliciting information on the respondents current health status and expectations of their heath in the future. The latter is particularly important because cancer risks in the survey won't result in mortality until several years in the future.

$ *Characterization of auto and stomach cancer risks and outcomes.* The respondents are provided with information on the general levels of risk for fatal automobile accidents and stomach cancers, and are informed that these risks may vary depending upon where one lives. The health outcomes are presented in detail, including how fatalities and automobile accidents differ with respect to timing.

$ *Familiarity with auto deaths and stomach cancer.* Because relative preferences for types of risk reduction may vary with experience of the health effects in question, the survey asks respondents about their familiarity with the risks in question.

$ *Choice description.* The basic framework for the choice question is presented to the respondents. Three simplified practice questions, to test for comprehension, are followed by three iterative choice questions. Subsequent choice questions are contingent upon prior responses in order to establish an interval over which the respondents preferred alternative changes. Indifference between alternative risk pairings lies in this interval.

- **$** *Risk perception and behavior questions.* Finally, the respondents are asked about their perceptions of risk and their behaviors that might affect the risks they face. These include driving habits, diet, etc.

Results from Cognitive Interviews

The current survey instrument has evolved substantially through a series of pretests, most of which have involved one-on-one cognitive interviews. The structure of each of these interviews was to first allow respondents to take the survey on their own and then to go back through the survey and verbally debrief each respondent

Though limited in number, these interviews revealed that respondents generally accept and understand the choice scenario, and exhibit a relatively even distribution of preferences. Some individuals preferred to reduce automobile accident risks for a variety of reasons for reasons that center around the timing of automobile mortality risk in the survey. Others preferred to reduce cancer risks in order to reduce the probability of pain and suffering associated with cancer.

Initial interviews based on a pencil-and-paper format had a number of inherent limitations (e.g., difficulties with skip patterns); however, these interviews did confirm the need to be very explicit in the survey about how latency works (i.e., reduced exposure this year implies reduced risk of death in X years). They also confirmed that choice framework needed to be as simple as possible – respondents were easily confused by information on two types of risks (car accident and cancer) for three locations (actual and 2 hypothetical).

Based in part on the lessons learned from these interviews, we turned to a computer-based interview format, using Knowledge Network's WebTV platform. In addition to this format change, a key difference in this instrument was how the choice questions were presented to respondents. In the earlier version, the two hypothetical locations were described in terms of *reductions* (rather than levels) of risk.

Although the respondents generally indicated a clear understanding of the latency period issue, several of the nine respondents had trouble with this choice format. Their verbal responses regarding which type of risk they viewed as worse were not consistent with their selection of area. Follow up choices which decreased the risk reduction in the selected area seemed to further exacerbate confusion with the format. For this reason, we revised the survey to present risks as in terms of levels.

While this change in framing the choice question did not completely eliminate confusion, it did appear to improve overall understanding of the choice. Most respondents made clear choices that were consistent with their verbal descriptions, although there were some inconsistences in the responses of some other three respondents. We have since further revised the survey to include a follow up questions which asks respondents to confirm their choice in order to reduce these inconsistencies.

Inconsistencies in responses to risk tradeoff questions have been noted in earlier work. Cropper and Krupnick (1991), for example, catalogs five types of inconsistent responses and provide some insights into how they may be treated in empirical analysis. It remains to be seen exactly how much this is an issue for a broader implementation of the survey, but such difficulties seem a worthwhile tradeoff to avoid many of the very difficult risk communication issues associated with direct elicitation of risk-dollar tradeoffs. By simplifying the cognitive difficulties involved in comprehending and weighing very small risks, we have been able to focus our efforts on communicating information on the attributes of the types of risks themselves, particularly the timing of the two fatal events. The cognitive interview results to date suggest that this strategy is successful and that respondents understand the essential tradeoff involved and accept the latency periods and choice framework.

References

Hammitt JK and Graham JD (1999). "Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability?" Journal of Risk and Uncertainty 18(1): 33-62.

Cropper, M.L. and A.J. Krupnick (1992). "The Effect of Information on Health Risk Valuations," *Journal of Risk and Uncertainty* 5(February): 29-47.

Magat, WA, Viscusi, WK, and Huber, J. (1996). "A Reference Lottery Metric for Valuing Health," *Management Science* 42:1118-1130.

Viscusi, W. Kip, Wesley A. Magat, and Joel Huber (1991). "Pricing Environmental Health Risks: Survey Assessments of risk-Risk and Risk-Dollar Trade-offs for Chronic Bronchitis," *Journal of Environmental Economics and Management,* 21: 32-51.

Eliciting Individual Beliefs About Mortality Risks
--Working Paper*--

PRESENTED BY:
GLENN W. HARRISON
UNIVERSITY OF SOUTH CAROLINA

CO-AUTHOR:
ELISABET RUTSTRÖM
UNIVERSITY OF SOUTH CAROLINA

*This is a working paper developed for the US Environmental Protection Agency National Center for Environmental Economics and National Center for Environmental Research's workshop, "Economic Valuation of Mortality Risk Reduction: Assessing the State of the Art for Policy Applications," held November 6-7, 2001 at the Holiday Inn Silver Spring Hotel in Silver Spring, MD.

Eliciting Individual Beliefs About Mortality Risks

by

Glenn W. Harrison and E. Elisabet Rutström ❖

DRAFT: August 3, 2001

NOT FOR DISTRIBUTION

Table of Contents

❖ Dewey J. Johnson Professor of Economics and Associate Professor of Economics, respectively, Department of Economics, The Moore School of Business, University of South Carolina (HARRISON@DARLA.BADM.SC.EDU and LISAR@DARLA.BADM.SC.EDU).

A. Summary

The monetary value of a statistical life (VSL) has been an important input into environmental cost-benefit analysis. Many environmental regulations entail some expectation of a reduction in premature mortality. This reduction in mortality constitutes a major benefit from environmental regulations. One important example is the retrospective EPA cost-benefit analysis of the provisions of the Clean Air Act, which concluded that mortality reduction benefits represented over $16 billion in present value terms.

The VSL can be inferred from labor market choices that workers make over alternative risky jobs. These calculations are conceptually straightforward, but in practice it has been hard to find ideal measures of the key variables needed to make the simple logic operational. Proxies for the subjectively perceived risk suffer from many problems. All existing studies appropriately caution that the empirical application of one proxy or another can often change the estimated value dramatically. This project examines several sources of uncertainty over the risk proxy, and evaluates the impact that better measures in the proxy have on VSL estimates. Improvements in the VSL methodology are of utmost importance for reliable assessment of health risks associated with the quality of the environment. The project seeks funding to improve one component of the overall VSL methodology: the risk proxy used. The project will therefore contribute towards the second focus area of this competition (research into individual and market valuation).

The overall goal is to develop methods for the reliable elicitation of the subjective expectations that individuals have concerning their risk of death from various causes. First, the project will develop better measures of the objective death risk than those currently available. This will involve developing an *"objective measure" of the risk of death for individuals that provides some degree of individual specificity*. That is, the death risk measure will be matched to the individual in terms of certain standard demographic characteristics, at least age, gender and race. This will utilize two data bases from the National Center for Health Statistics, and the estimation of multinomial logit models. This stage of the project will also measure latent risk, and will focus on the entire distribution of risk, and not just some simple descriptive measure such as the mean. Second, the project will *develop a method for estimating differences between subjectively perceived risks and objective risks*, and will apply this method on a convenience subject sample. This will employ techniques from

Experimental Economics, implementing a standard scoring rule elicitation mechanism. Third, it will *develop a method for estimating calibration factors* for these differences which can be used to calibrate the objective risk measures to better reflect subjectively perceived risk, and to again apply this method on a convenience subject sample. This stage of the project will integrate the results of the first two stages. The project will also make it possible to assess the sensitivity of these differences to variations in the method of getting individual reports, such as the presence of financial incentives and the method used to communicate risk. The envisioned use of the methodology developed is to provide better proxies for subjectively perceived risk than the currently used objective industry mean risks.

The value of improved precision in the risk proxies used to construct the VSL for environmental cost-benefit analysis should be obvious. Reductions in mortality rates are important aspects of the benefit side of the equation, and good estimates of the VSL are necessary for the valuation of these reductions. The estimation of calibration factors, in particular, will make future estimations of risk proxies more accurate.

B. Project Description

The monetary value of a statistical life (VSL) has been an important input into environmental cost-benefit analysis. Many environmental regulations entail some expectation of a reduction in premature mortality. This reduction in mortality constitutes a major benefit from environmental policies. One important example is the retrospective USEPA [1997] cost-benefit analysis of the provisions of the Clean Air Act, which concluded that mortality reduction benefits represented over $16 billion in present value terms. Uncertainty over the VSL translates directly into uncertainty over the benefits of environmental policies, such as those embodied in the Clean Air Act (e.g., see USEPA [1997; Figure I-2, p. I-21]).

VSL estimates have been inferred from labor market choices that workers make over alternative risky jobs, or from the prices consumers have paid for goods with negative environmental or health effects. In these cases the wages or prices are assumed to include a risk premium, reflecting the VSL, and estimates of these risk premia rely on the use of proxies for risk as explanatory variables. These calculations are conceptually straightforward, but in practice it has been hard to find ideal measures of the key explanatory variables needed to make the simple logic operational.[1] Problems are encountered when constructing the proxies for the subjectively perceived risk. The focus on individual subjective expectations is appropriate since this is the basis upon which individuals make their private decisions.

Lacking measures of the subjective risk beliefs, investigators have resorted to the use of objective measures, but even these are problematic. Excessive aggregation of the classifications of risky activities and the lack of measures of latent risk are two examples of such problems with the objective measures used. These problems are often caused by the lack of completeness and accuracy

---

[1] The basic idea underlying the VSL calculation is that the observed labor market choices of workers will reflect their willingness to accept risks, just as it reflects other aspects of the job. All other things being equal, a worker that is risk averse would require a higher wage to work in a job that is riskier than some alternative job. Thus one would expect that the different wages observed for different jobs should be positively correlated with the riskiness of the job. If so, one can back out a measure of the extra dollars per year a typical worker requires to accept a job that has an extra 1-in-10,000 chance of being associated with a fatal injury. If this wage premium, or risk premium, is $500, then we can multiply by 10,000 to ascertain the value that is *implicitly* placed on a *statistically* certain death. In this case the value would be $500 × 10,000 = $5 million, which happens to be a good approximation of the ball-park of many of the estimates (e.g., Viscusi [1993; Table 2]).

of the field data used. All existing studies appropriately caution that the empirical application of one proxy or another proxy can often change the estimated value dramatically. This project examines several sources of the unreliability of these risk proxies. A better understanding of these sources will greatly improve our ability to construct more accurate and reliable measures of VSL. The project seeks funding to improve one specific component of the overall VSL methodology: the risk proxy used. The project will therefore contribute towards the second focus area of this competition (research into individual and market valuation).

The overall goal is to develop methods for the reliable elicitation of the subjective expectations that individuals have concerning their risk of death from various causes. Our project will integrate valuation and decision-making approaches for environmental policy, since the VSL plays such an integral role in environmental policy analysis.

Specifically, our project will (i) find better measures of the objective death risk, and evaluate the impact this has on VSL estimates; (ii) develop a method for estimating differences between subjectively perceived risks and objective risks, and apply this method on a convenience subject sample; and (iii) develop a method for estimating calibration factors for these differences which can be used to correct objective risk measures to better reflect subjectively perceived risk, and to again apply this method on a convenience subject sample. The project will also make it possible to assess the sensitivity of these differences to variations in the method of getting individual reports, such as the presence of salient incentives and the method used to communicate risk. The following three major steps will lead to the fulfillment of the stated goal:

1. The first step is to develop an *"objective measure" of the risk of death for individuals that provides some degree of individual specificity*. That is, we would like to have some "true" death risk measure that can be matched to the individual in terms of certain standard demographic characteristics, at least age, gender and race. This will be used as a comparison for the subjective beliefs elicited in stages 2 and 3. Thus, measures of bias in the subjective responses can be calculated.
2. The second step is to develop a laboratory procedure which provides individuals with clear *financial incentives to state their beliefs*. It has been shown in a number of previous studies that financial incentives have the power to dominate and overcome many sources of

response variations and biases (e.g., Cummings and Harrison [1994], Cummings, Harrison and Rutström [1995], Cummings, Elliott, Harrison and Murphy [1997], and Neill et al. [1994]). This procedure should enable us to elicit the distribution of their beliefs, as well as certain summary statistics of that distribution such as the mean. The instrument will be computerized and web-based, allowing it to be transported into the field with relative ease by future users.

3. The third step is to examine the *effect of alternative methods of "risk communication"* that have been proposed in the literature. It is well-known that there are differences in elicited beliefs when different visual and linguistic devices are used to describe the risk in question. Do the differences observed in studies on risk diminish or become more predictable when we use financial incentives to elicit the beliefs and tailor the risk to the individual more?

These steps can be performed by combining tools from existing studies. These are discussed further in section 2, below.

Section 1 provides further motivation for the concern over the risk proxies currently used, drawing from the VSL literature. Section 2 explains the three steps in this project. Section 3 reviews the proposed budget, and section 4 reviews qualifications and facilities.

1. Why do we need better proxies for the risk of death?

Four issues are of some concern in the existing VSL literature, and are addressed by our proposal to some degree: excessively aggregated measures of risk, the treatment of proxies for risk which are uncertain as though they were exact, the focus on contemporaneous risks that occur "on the job," and the likelihood of differences between objective expectations of risk and subjective expectations of the same risk. There are many other issues of concern in the VSL literature, but we limit our focus to these four.

Aggregation. Many studies relate the wages of individual workers with risk measures calculated for the industry that those workers worked in, with no accounting for the occupation of the individual worker.[2] Thus an accountant working in the logging industry is assumed to face the

[2] Krupnick, Alberini, Cropper and Simon [1999; p.3] correctly note that the "main shortcoming of labor market studies is that they measure compensation received by prime-age men for immediate reductions

same risk as a lumberjack in that industry. This issue has been well-known, particularly since Dillingham [1985]. Leigh [1995] and Scotton and Taylor [2000] are the most recent studies to take these factors into account. The approach adopted here allows us to develop a proxy variable for the risk of death faced by workers varying by many characteristics in addition to industry and occupation. This proxy variable will be used as the objective measure of the risk faced by the individual, as described later.

Uncertainty. The approach adopted here also allows for uncertainty in the risk proxy, which is typically ignored by assuming that each worker in a given industry faces a fixed average risk with no variation.[3] The standard approach can be viewed as akin to a statistical imputation, in which some deterministic value is assumed to be the true value for the purposes of analysis. Whenever there is some uncertainty about the imputed value, due to the very fact that it is imputed, the analysis that assumes this uncertainty away could be led to make false inferences. There is some inevitable measurement error in the imputed proxy which should be taken into account. When the imputed proxy is itself just the mean of some distribution, this uncertainty is transparent. Our objective measure of risk takes this uncertainty into account.

Latency. Apart from the problem of allowing for the occupation and other characteristics of the worker within an industry, there is a problem with the proxy measures only capturing contemporaneous risks, as emphasized by Johannesson and Johansson [1996], Johannesson, Johansson and Löfgren [1997], and Krupnick, Alberini, Cropper and Simon [1999], amongst others. That is, most of the risk measures capture the effects of fatal accidents occurring *on the job*, and ignore the *latent* risks of being *in one industry or occupation*. If the worker is assumed to take into account the contemporaneous risks when agreeing to work for a given risk premium, then it is

---

in risk of death." Although some of the criticisms implied by this comment can be dealt with using simple extensions of the labor market studies (e.g., by including female workers in the sample), the reason that prime-age risk measures are used is that virtually all studies evaluate the VSL at the sample mean. It would be a simple matter to stratify the sample so as to generate VSL estimates for younger or older workers. But such stratification effectively requires that there be some disaggregation in the underlying risk measure in terms of age and other factors, such as proposed to be undertaken here.

[3] This uncertainty in the risk *proxy* itself is quite different from the *subjective* uncertainty that the worker has when making an estimate of the riskiness of a job. The latter should reflect the former, to the extent that there is some objective uncertainty in the measurement of risk, but is distinct.

logical that the worker also takes into account risks that could materialize later in life. The empirical importance of this aspect of the possible risk measure is apparent when one notes that the vast majority of people die from causes which are *not* due to contemporaneous on-the-job accidents. For example, in 1993 about 2.2 million Americans died, and only 5,000 or so could be classified as dying from occupational injuries.[4]

Subjective Beliefs. The correct conceptual variable in these wage regressions is the subjective expectation of risk, not the objective expectation. This point was stressed by Gerking, de Haan and Schulze [1988] and Gegax, Gerking and Schulze [1991], who directly undertook a field survey of individuals to elicit their subjective beliefs.[5] Their survey was hypothetical, in the sense that it provided the subjects no financial incentives to respond accurately. Nonetheless, it provided a rich and important alternative to the objective measures. Their subjective belief measures of fatal risk resulted in VSL estimates in Gegax, Gerking and Schulze [1991] that were "substantially lower than those obtained using an industry average risk measure." (p. 595). Indeed, for white-collar and non-unionized workers they find that the VSL is not statistically different from zero.

We undertake controlled laboratory experiments in which subjects are provided financial incentives to reveal the distribution of their subjective beliefs about objective risks of death. We then use these elicited responses to design calibration factors that allow objective measures of risk to be adjusted to better reflect subjective risks. These calibration factors also provide some guidance to policy-makers who might choose to substitute objective measures of risk when determining which VSL to use for policy. Blackburn, Harrison and Rutström [1994] and Harrison et al. [1999] review the use of experimental methods in the construction of calibration factors for field data.

---

[4] Source: National Center for Health Statistics, National Mortality Followback Survey, Provisional Data -- Public Use Data File, 1993. The standard definition of an occupational fatality is when the death is coded as resulting from an "external cause" and the item "injury at work" is marked on the death certificate.

[5] The survey was undertaken using a questionnaire that was mailed to subjects. A copy of the survey instrument is available at HTTP://DMSWEB.BADM.SC.EDU/GLENN/GEGAXGERKINGSCHULZE_SURVEYINSTRUMENT.PDF.

2. This Project

A. Objective Measures of the Risk of Death

The first step of our project involves the construction of better and more reliable measures of the objective risks from health hazards associated with the environment.

Two databases generated by the National Center for Health Statistics [1997] [1998] provide the opportunity to construct more detailed proxies for risk. The first is the Multiple Causes of Death file linked to the 1986-1994 National Health Interview Surveys (NHIS), and the second is the 1993 National Mortality Followback Survey (NMFS). Each provides detailed data on the nature of the death, as well as relatively rich demographics on the decedent.

The NHIS is an annual survey undertaken by the U.S. Department of Health and Human Services for many years. In 1995 the National Center for Health Statistics (NCHS) merged the NHIS data from 1986 through 1994 with comprehensive data from national death certificates. The NCHS attempted to search the national death certificate database and identify the respondents to the NHIS surveys between 1986 and 1994. In many cases they were successful, since they had access to necessary private, identifying information on the respondent (e.g., Social Security Number, name). The result is a database of 430,000 respondents, 37,706 of whom died during the 1986-1995 time period. Table 1 shows how these data relate to the original NHIS survey. For each death, the certificate provided information to identify up to 282 "standard" causes of death. The original NHIS survey asked many questions about the characteristics of the individual, including 98 standard occupations and 98 standard industries.[6] The usual socio-demographic variables, such as age, gender, race, marital status, and education level, are also included.

The 1993 NMFS is the sixth mortality followback survey conducted by NCHS. The NMFS focuses on a nationally representative sample of 19,018 individuals aged 15 years and over who died in 1993. Forty-nine of the 50 State vital registration areas granted approval to sample their death certificates.[7] Data for the 1993 NMFS are therefore representative of individuals aged 15 years or

---

[6] In fact the survey contains more detailed information, but there are 98 convenient "re-codes" of the most detailed information which provide sufficient detail for present purposes.

[7] The independent vital registration areas of the District of Columbia and New York City also gave permission to sample their death certificates. South Dakota declined to participate in the NMFS due to state law restricting the use of death certificate information.

older who were residing and died in the United States in 1993, excluding South Dakota. The data from the NMFS were collected from death certificates[8] and proxy interviews[9]. In addition to detailed information on the circumstances of death, the NMFS collected information on the usual detailed industry and occupation of the decedent, as well as the usual socio-demographic characteristics. Table 2 shows some of the data from the NMFS in terms of major causes of death (the underlying database contains much more detailed information on causes of death).

The NHIS and NMFS data allow the development of multinomial logit models to predict the risk of death, by cause-of-death, as a function of individual characteristics of the NHIS and NMFS samples. Since these characteristics include industry and occupation, the risk proxies developed with this methodology generate the same detail as Scotton and Taylor [2000]. However, the addition of other individual demographic characteristics and job-related characteristics allows the proxy to be a better reflection of the expected death risks for the individual.

The statistical model estimated using the NHIS and NMFS data can be used to predict the *expected* risk of death for individuals based on their demographic characteristics, industry, and occupation, in the main VSL estimation sample. However, this approach also allows one to predict

---

[8] A sample of 22,957 death certificates was drawn from the 1993 Current Mortality Sample (CMS) which formed the sampling frame. The CMS is a 10-percent systematic sample of death certificates received at NCHS from state vital statistics offices. The sample was selected by broad age groups, two racial groups, and gender within 12 causes of death (suicide; homicide; injuries to motor vehicle drivers, pedestrians, and motorcycle owner operators; other motor vehicle injuries; non-motor vehicle injuries; HIV; cancer; COPD; heart disease; alcohol abuse; drug abuse; and all other causes). In order to produce more robust analysis, black decedents, certain causes of death, and certain age groups were over sampled (45.5 percent of all cases). For example, a little more than 1/3 of the sample represent decedents less than 35 years of age, about 4 percent represent decedents over 100 years of age, almost 31 percent of the sample represents black decedents, 67 percent represent male decedents, and a little less than 44 percent represent persons who died of external (homicide, suicide, accidental injury) causes.

[9] Decedent's next-of-kin (respondents), identified on the death certificate as having provided information, were initially contacted by personalized letter and asked to participate in the survey. In cases where no next-of-kin information was available from the death certificate, letters were sent to funeral directors requesting contact information for the next-of-kin. The U.S. Bureau of the Census, acting as data collection agent, subsequently contacted the next-of-kin respondents by telephone (about 65 percent of the sample) or in-person to obtain the necessary data. When no next of kin was identified, or they could not be located, another person familiar with the decedent's life history was requested to provide the necessary information. An overall 83 percent response rate was achieved.

the *distribution* of risk facing the individual, apart from just predicting the *mean* risk facing the individual. Thus it will be possible to see how objective uncertainty over the risk of death translates into uncertainty over the VSL. Previous studies have implicitly assumed that the proxy value generated was known with zero measurement error, just as precisely as we know the age or gender of the worker. In fact, it is apparent that there may be considerable error in this proxy variable, and that some of it may be correlated with variables of interest in the estimation stage (e.g., as stressed by Leigh [1995]).

Thus the risk proxy variable developed in this stage of the project has three advantages over previous approaches: it will be specific to individual characteristics, it will reflect the uncertainty in the process generating the proxy, and it includes all causes of death (and not just the deaths that occur on the job). For the purposes of environmental decision making this is obviously more appropriate. Thus we believe that the results from this first stage of the project will be of great value in improving VSL estimates, quite apart from the methodological development of instruments to elicit subjective beliefs considered in the second and third stages of the project.

B. Eliciting Beliefs About the Risk of Death

The true risk of death for an individual is a stochastic variable. From the econometric analysis undertaken in the first part of the project we will obtain an estimate of the empirical *distribution* of that risk for an individual. That estimate can be conditioned on any of the explanatory variables in our model, but obviously (and minimally) on age, gender and race. Since we know the empirical distribution, we can also define the mean risk, the median risk, and confidence intervals on that risk.

Ignoring for the moment how we communicate these risks to the subject, since that is an issue we address below in section C, the second stage of our project develops laboratory instruments to elicit the subject beliefs that individuals have about these risks. The main contribution here is to use financial incentives and appropriate "scoring rules" to motivate subjects to reveal their beliefs in this task.[10]

[10] The quadratic scoring rule has been used by McKelvey and Page [1990] and Grether [1992], and has appeared in numerous unpublished experimental studies.

We utilize a quadratic scoring rule procedure. Let MR denote the mortality risk investigated. We will elicit not just the mean of an individual's subjective beliefs over MR, but an approximation to the entire belief distribution. We therefore elicit the probability distribution over the entire range of risk factors, MR = [0.01,... 1.00]. We will convert the true value to discrete intervals, such as "less than 0.1, between 0.1 and below 0.2, etc.," since this will make it easier for the subjects to understand the task without any essential loss in accuracy. Following the standard textbook exposition in Davis and Holt [1993; p.465ff.], let P denote the individual's *subjective probability* that a person "just like them" faces a mortality risk corresponding to one specific such discrete risk intervals. This P is what we want to elicit truthfully. Let R denote the same individual's *report* of that probability. Then let I denote the *outcome*, which in this case is equal to 1 only for the discrete risk interval containing the true risk, and 0 otherwise. The scoring rule then provides payoffs equal to $1 - (R-I)^2$, which is equal to $1-R^2$ if $I=0$ and $2R-R^2$ if $I=1$. Since each individual attaches true probability P to the outcome I being true, the *expected* payoff to a risk-neutral subject would be

$$P\,(2R - R^2) + (1-P)\,(1-R^2).$$

This payoff rule gives the subject an incentive to report the value P, since expected payoff is maximized when $R = P$. In the jargon of decision theory, such a rule is said to be a "proper" scoring rule since the subject maximizes expected payoff by directly reporting their true probability; as noted by Davis and Holt [1993; p.467, f.25], there are many other scoring rules that are not proper but for which one can still infer the underlying probability P from the report R.[11]

The task to be presented to subjects will be a simpler version of this formal logic. Assume that we have defined what characteristics the subject should use in determining if someone is "just like them." For example, it could be someone in the same age bracket, the same gender, and the same racial group. In the section called "Protocol" we review the essentials of the experimental protocol that will be used to elicit risks.

We propose an experimental design with four different mechanisms for eliciting the beliefs and several different risk communication methods. The four belief elicitation methods are:

1. Standard Scoring rule with standard payoffs

[11] In the jargon of mechanism design theory, *proper* rules are *direct* revelation mechanisms.

2. Standard Scoring rule with significantly higher payoffs

3. Standard Scoring rule without financial incentives

4. Simple, discrete choice elicitation method with standard payoffs.

The purpose of this four-way design is to test the impact of financial incentives on the willingness and ability of subjects to report their subjective beliefs using the scoring rule. Our hypothesis is that the treatment with significantly higher payoffs will result in more accurate reports, resulting in reports with smaller variance. Due to the potential cognitive complexity of the scoring rule, we also evaluate a behaviorally simpler method: the discrete choice rule, which simply asks subjects to report a value for the risk, paying them if they are sufficiently close to the truth, but not otherwise.

Subjective valuations have been elicited in experimental settings in many previous studies (for example, Rutström [1998], Cummings, Harrison and Rutström [1995], and Fox, Shogren, Hayes, and Kliebenstein [1996]). Similarly, some studies have estimated calibration factors in subjective valuations settings, such as Fox, Shogren, Hayes, and Kliebenstein [1998], and Blackburn, Harrison, and Rutström [1994]. For a review of the importance of financial salience, see the review by Harrison and Rutström [2000]. Recently, some experimental studies have employed the proper scoring rule to elicit subjective beliefs; see, for example, McDaniel and Rutström [2000] and Morrison and Rutström [1999].

C. Evaluating Alternative Risk Communication Methods

There are many ways to present probabilities and risks to subjects. The subjects in our experiments need to have a firm understanding of the notion of "probability," so that their responses can be viewed as based on a meaningful use of the concept of risk. Even though we are concerned with the task of *eliciting* probability, we need to ensure that the notion of probability is *communicated* to the subject so that we have some confidence that their responses are meaningful.

The available literature suggests that subjects can be sensitive to different methods of explaining risk. The "language" of risk communication includes verbal methods as well as visual methods. In turn, there are many variants on each of these approaches. Verbal methods can be analogical as well as numerical, as illustrated by Calman and Royston [1997]. Similarly, visual methods can utilize "risk ladders" or "pie charts" or "dots," as illustrated by Loomis and du Vair

[1993] and Hammitt and Graham [1999]. A review of the literature by Hammitt and Gordon [1999; Table 1, p.37/8] indicates that the mode of risk communication has varied widely, and it is plausible that these differences can account for much of the variation in elicited values. The current trend, illustrated by Krupnick, Alberini, Cropper and Simon [1999; p.8], is to use visual grids with no more than 1,000 cells, on the grounds that using more cells can hamper cognition.

Recognition of the importance of future risks of illness or death has been considered explicitly by Krupnick, Alberini, Cropper and Simon [1999], who develop instruments designed to facilitate the measurement of future risk. One of their instruments presents the "commodity" to be valued in terms of risk reductions, and the other instrument presents the "commodity" in terms of life expectancy changes. Given our concern with latent risks, an evaluation of the effects of their instruments will also be undertaken.

We propose reviewing the literature and employing a range of alternative risk communication methods, so that we can determine which work best. The five different communication method designs which we will consider are: verbal analogical, verbal numerical, visual ladders, visual pie charts, and visual dots. We also want to consider one treatment in which we combine the best "verbal" method with the best "visual" method, so we would have six communication method treatments in all. The "Protocol" section below provides some specific examples of these methods. It would not be surprising to see some methods of risk communication work better for some groups, and others work better for other groups.

Our objective is not to come up with new methods of risk communication, so much as to allow for the possibility that the use of different methods can affect responses. It is to be hoped that differences in responses found in the existing literature are mitigated when we provide financial incentives and individualized questions, but we do not want to pre-judge the behavioral issue by restricting ourselves to one specific risk communication language.

3. Policy Relevance and Outreach

The usefulness of VSL estimates for environmental policy is apparent, and well illustrated by the retrospective cost-benefit analysis of the Clean Air Act undertaken by the USEPA [1997]. Many citations to other policy applications of VSL estimates can be found in the literature; Mrozek

and Taylor [1999; p.3], for example, cite prominent applications by the FDA, OMB and DOT.

Our project will improve the quality of the VSL estimates by providing information on objective risks that is more individual-specific than previously developed, by including latent risks, by providing information on the uncertainty over those risks, and by developing measures that allow objective risks to be calibrated to better approximate subjectively perceived risks. Our project entails several methodological improvements in the standard approach, by using complementary data sources such as the NMFS and experimental methods. We will also provide new VSL estimates using our improved risk proxy measures.

Our data will be made available through the web for others to use. We will provide all statistical software and data output. In particular, we will provide the "new and improved" risk measures and calibration factors in the form of STATA, SAS, and SPSS data files, to allow researchers using those packages to easily test the effect of our suggested measures.

Our findings will also be submitted for publication to top ranked academic journals (both general ones and specialty ones), such as the *Quarterly Journal of Economics*, *American Economic Review*, and the *Journal of Environmental Economics and Management*. Working papers will be presented at two or more academic conferences; one conference would be for environmental economists (such as the Second World Congress of Environmental and Resource Economists), and one would be for experimental economists (such as the International Meetings of the Economic Science Association). Working papers will also be published on the web, at well frequented sites such as EconWPA (HTTP://ECONWPA.WUSTL.EDU/WPAWELCOME.HTML).

4. Protocol

In order to elicit subjects belief distribution over the true mortality risk, we employ a proper scoring rule, as discussed in the main project description.

Consider the case in which we are finding our what the subject's beliefs are that 100 people "just like them" will eventually die from Heart Disease. Each subject will be presented with a Report Card, consisting of ten intervals of numbers, ranging from(0-10) to (91-100), corresponding to possible beliefs. Table 4 illustrates this Report Card. Subjects will also have been given a Payoff Table, such as the one shown in Table 3, which illustrates how their reports translate

into payoffs. The Payoff Table shows ten different reports that the subjects could make, corresponding to different probabilities, and two possible states of nature. In one state the TRUE RISK *is* found in a particular interval; in the other state, the TRUE RISK *is not* found in that interval. There is a payoff associated with each report contingent on either of these states being true. Every subject would be asked to make a report for *each* of the ten intervals of numbers on their Report Card, corresponding to their belief about the likelihood that the TRUE RISK is in that interval, and they would be paid for all ten. For example, if the TRUE RISK for this subject was 31 (out of 100), and the subject made a report of 0.9 for the interval (51-60), he would receive $0.19 for that interval since the TRUE RISK is not a number between 51 and 60. If this subject made a report of 0 for the interval (31-40), he would receive $0.00 for that interval since the TRUE RISK is 31, etc.

This task will be initially presented to subjects using a series of simplified training experiment, such as estimating the number of black marbles in a large jar of black and white marbles. The training experiments will primarily serve the purpose of explaining the task to the subjects and helping them understand the incentive compatible nature of the incentive mechanism, so as to maximize the propensity for truthfulness and accuracy. An additional trainer will ask subjects to elicit their beliefs about the *height* of someone just like them, allowing them to train in a task with a similar context to the one in the task we want to observe, i.e. the mortality risk. Subjects will receive some payments in these training sessions as well.

In addition to these elicitation mechanisms, we will also employ a number of different ways of illustrating or communicating the risk. The subjects in our experiments need to have a firm understanding of the notion of "probability," so that their responses can be viewed as based on a meaningful use of the concept of risk. Even though we are concerned with the task of *eliciting* probability, we need to ensure that the notion of probability is *communicated* to the subject so that we have some confidence that their responses are meaningful. It has been found in other research that some visual aids facilitate in subjects processing information, and we propose to test whether this is the case for risk elicitation. The five different communication method designs are: verbal analogical, verbal numerical, visual ladders, visual pie charts, and visual dots. We also want to consider one treatment in which we combine the best "verbal" method with the best "visual" method, so we would have six communication method treatments in all.

An example of the verbal analogical scale is provided by Calman and Royston [1997; Table 4], using a distance analogue. For risks of 1 in 1, 1 in 10, 1 in 100, 1 in 1000, for example, the distance containing one "risk stick" 1 foot in length is 1 foot, 10 feet, 100 feet, and 1,000 feet, respectively. The visual dots method is employed by Krupnick, Alberini, Cropper and Simon [1999], and provides a graphic image to complement the direct fractional, numerical representation of probability. To implement this risk communication method we would provide an array of graphic images to match each of the 10 intervals show in the Report Card in Table 4. Visual ladders have been used in previous VSL research by Gerking, de Haan and Schulze [1988] and Gegax, Gerking and Schulze [1991]; we include the risk ladder from their survey instrument at the end of this section. It would be possible to adapt this risk ladder to better reflect the characteristics of the individual respondent, by first eliciting some basic socio-demographic information about the individual and then generating a real-time ladder appropriate for that individual. Or, in field applications, elicit the occupation of the individual and then use that information to generate a risk ladder showing germane levels of risk for that individual.

Table 1: Number of Deaths by Survey Year and Year of Death in the NHIS Database

	NHIS Survey Year									
Actual Year of Death	1986	1987	1988	1989	1990	1991	1992	1993	1994	1986/94
1986	161	-	-	-	-	-	-	-	-	
1987	364	351	-	-	-	-	-	-	-	
1988	432	757	345	-	-	-	-	-	-	
1989	439	846	793	347	-	-	-	-	-	
1990	460	832	834	733	326	-	-	-	-	
1991	428	903	850	779	758	321	-	-	-	
1992	472	890	828	844	837	740	372	-	-	
1993	519	915	991	913	886	826	849	284	-	
1994	493	1,015	954	856	887	875	872	686	288	
1995	471	1,046	971	938	868	919	878	743	721	
Total	4,239	7,555	6,566	5,410	4562	3,681	2,971	1,713	1,009	37706

Table 2: Number of Deaths by Cause of Death in the NMFS Database

Sample size, estimated number and percent of deaths to persons aged 15 years or older residing in the United States (excluding South Dakota) in 1993, and standard errors of percents, by cause-of-death sampling criteria

Cause-of-death sampling criteria	Responding sample size	Estimates		Standard error of percent
		Number	Percent	
All causes	19,018	2,215,440	100.0	0.00
Suicide	1,616	31,623	1.4	0.05
Homicide	1,493	24,551	1.1	0.01
Motor vehicle drivers, motorcycle operators, pedestrians	1,766	23,179	1.0	0.01
Other motor vehicle injuries	1,055	14,781	0.7	0.01
Other non-motor vehicle injuries	2,134	44,770	2.0	0.05
Alcohol abuse	1,386	18,157	0.8	0.01
Drug abuse	180	3,492	0.2	0.04
HIV	2,011	39,693	1.8	0.06
Cancer	1,655	526,432	23.8	0.16
COPD	1,069	101,806	4.6	0.07
Heart disease	1,949	730,492	33.0	0.22
All other causes	2,704	656,463	29.6	0.24

Table 3: Illustrative Payoff Table

Choose your REPORT from the following:	If the TRUE RISK **IS** in the interval you will get:	If the TRUE RISK is **NOT** in the interval you will get:
1.0	$1	$0
0.9	$0.99	$0.19
0.8	$0.96	$0.36
0.7	$0.91	$0.51
0.6	$0.84	$0.64
0.5	$0.75	$0.75
0.4	$0.64	$0.84
0.3	$0.51	$0.91
0.2	$0.36	$0.96
0.1	$0.19	$0.99
0	$0	$1

Table 4: Illustrative Report Card

INTERVAL	YOUR REPORT	YOUR PAYMENT (PLEASE DO **NOT** WRITE IN THIS COLUMN)
(0-10)		
(11-20)		
(21-30)		
(31-40)		
(41-50)		
(51-60)		
(61-70)		
(71-80)		
(81-90)		
(91-100)		
		Total: _____

D. References

Blackburn, McKinley; Harrison, Glenn W., and Rutström, E. Elisabet, "Statistical Bias Functions and Informative Hypothetical Surveys," *American Journal of Agricultural Economics*, 76(5), December 1994, 1084-1088.

Calman, Kenneth C., and Royston, Geoff, "Risk Language and Dialects," *British Medical Journal*, 315, 1997, 939-942; available as HTTP://DMSWEB.BADM.SC.EDU/GLENN/RISKLANGUAGE.PDF.

Cummings, Ronald G.; Elliott, Steven; Harrison, Glenn W., and Murphy, James, "Are Hypothetical Referenda Incentive Compatible?" *Journal of Political Economy*, 105(3), June 1997, 609-621.

Cummings, Ronald G., and Harrison, Glenn W., "Was the *Ohio* Court Well Informed in Their Assessment of the Accuracy of the Contingent Valuation Method?," *Natural Resources Journal*, 34(1), Winter 1994, 1-36.

Cummings, Ronald G.; Harrison, Glenn W., and Rutström, E. Elisabet, "Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible?" *American Economic Review*, 85(1), March 1995, 260-266.

Davis, Douglas D., and Holt, Charles A., *Experimental Economics* (Princeton: Princeton University Press, 1993.

Dillingham, Alan E., "The Influence of Risk Variable Definition on Value-of-Life Estimates," *Economic Inquiry*, XXIV, April 1985, 277-294.

Fox, John A., Shogren, Jason F., Hayes, Dermot J., and Kliebenstein, James B., "CVM-X: Calibrating Contingent Values with Experimental Auction Markets," *American Journal of Agricultural Economics*, 80, 1998, 455-465.

Fox, John A., Shogren, Jason F., Hayes, Dermot J., and Kliebenstein, James B., "Experimental Methods in Consumer Preference Studies." *Journal of Food Distribution Research*, July 1996, 1-7.

Hammitt, James K., and Graham, James D., "Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability?" *Journal of Risk and Uncertainty*, 18(1), 1999, 33-62.

Garen, John, "Compensating Wage Differentials and the Endogeneity of Job Riskiness," *Review of Economics and Statistics*, 70, 1988, 9-16.

Gegax, Douglas; Gerking, Shelby, and Schulze, William, "Perceived Risk and the Marginal Value of Safety," *The Review of Economics & Statistics*, 73, 1991, 589-596.

Gerking, Shelby; de Haan, Menno, and Schulze, William, "The Marginal Value of Job Safety: A Contingent Value Study," *Journal of Risk & Uncertainty*, 1, 1988, 185-199.

Grether, David M., "Testing Bayes Rule and the Representativeness Heuristic: Some Experimental Evidence," *Journal of Economic Behavior and Organization*, 1992, 17, 31-57.

Harrison, Glenn W., "Theory and Misbehavior of First-Price Auctions: Reply," *American Economic Review*, 82, December 1992, 1426-1443.

Harrison, Glenn W., "Expected Utility Theory and the Experimentalists," *Empirical Economics*, 19(2), 1994, 223-253; reprinted in J.D. Hey (ed.), *Experimental Economics* (Heidelberg: Physica-Verlag, 1994).

Harrison, Glenn W.; Beekman, Robert L.; Brown, Lloyd B.; Clements, Leianne A.; McDaniel, Tanga M.; Odom, Sherry L. and Williams, Melonie, "Environmental Damage Assessment With Hypothetical Surveys: The Calibration Approach," in M. Boman, R. Brännlund and B. Kriström (eds.), *Topics in Environmental Economics* (Amsterdam: Kluwer Academic Press, 1999).

Harrison, Glenn W., and Rutström, E. Elisabet, "Experimental Evidence of Hypothetical Bias in Value Elicitation Methods", forthcoming in C.R. Plott and V.L. Smith (eds.), *Handbook of Experimental Economics Results,* 2000.

Johannesson, Magnus, and Johansson, Per-Olov, "To Be, or Not to Be, That is the Question: An Empirical Study of the WTP for an Increased Life Expectancy at an Advanced Age," *Journal of Risk & Uncertainty*, 13, 1996, 163-174.

Johannesson, Magnus; Johansson, Per-Olov, and Löfgren, Karl-Gustaf, "On the Value of Changes in Life Expectancy: Blips Versus Parametric Changes," *Journal of Risk & Uncertainty*, 14, 1997, 221-239.

Krupnick, Alan; Alberini, Anna; Cropper, Maureen, and Simon, Nathalie, "Mortality Risk Valuation for Environmental Policy," *Discussion Paper 99-47*, Resources For the Future, 1999; available as HTTP://WWW.RFF.ORG/CFDOCS/DISC_PAPERS/PDF_FILES/9947.PDF.

Leigh, J. Paul, "Compensating Wages, Value of a Statistical Life, and Inter-industry Differentials," *Journal of Environmental Economics & Management*, 28(1), 1995, 83-97.

Loomis, John B., and du Vair, Pierre, "Evaluating the Effect of Alternative Risk Communication Devices on Willingness to Pay: Results from a Dichotomous Choice Contingent Valuation Experiment," *Land Economics*, 69, 1993, 287-298.

McDaniel, Tanga M., and Rutström, E. Elisabet, "Decision Making Costs and Problem Solving Performance," *Unpublished Manuscript*, Department of Economics, University of South Carolina, December 2000.

McKelvey, Richard D., and Page, Talbot, "Public and Private Information: An Experimental Study of Information Pooling," *Econometrica*, 58, 1990, 1321-1339.

Miller, Ted R.; Cohen, Mark A., and Wiersema, Brian, *Victim Costs and Consequences: A New Look* (Washington, DC: National Institute of Justice, February 1996); available as

HTTP://WWW.NCJRS.ORG/PDFFILES/VICTCOST.PDF.

Moore, Michael J., and Viscusi, W. Kip, *Compensation Mechanisms for Job Risks* (Princeton: Princeton University Press, 1990).

Morrison, William G., and Rutström, E. Elisabet, "On Reciprocity and Psychological Equilibria in an Investment Experiment," *Unpublished Manuscript*, Department of Economics, University of South Carolina, September 1999.

Mrozek, Janisz R., and Taylor, Laura O., "What Determines the Value of Life? A Meta Analysis," *Working Paper*, Department of Economics, Andrew Young School of Policy Studies, Georgia State University, 1999; available as HTTP://DMSWEB.BADM.SC.EDU/GLENN/MROZEKANDTAYLOR.PDF.

National Center for Health Statistics, *National Health Interview Survey -- Multiple Cause of Death Public Use Data File, 1986-1994 Survey Years*. National Center for Health Statistics, Division of Health Interview Statistics, Hyattsville, MD; 1997.

National Center for Health Statistics, *National Mortality Followback Survey -- Provisional Data, 1993. Public Use Data File Documentation*. National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Maryland; 1998.

Neill, Helen R.; Cummings, Ronald G.; Ganderton, Philip T.; Harrison, Glenn W., and McGuckin, Thomas, "Hypothetical Surveys and Real Economic Commitments," *Land Economics*, 70(2), May 1994, 145-154.

Rutström, E. Elisabet, "Home-Grown Values and the Design of Incentive Compatible Auctions," *International Journal of Game Theory*, 27(3), 1998, 427-441.

Scotton, Carol R., and Taylor, Laura O., "New Evidence From the Labor Markets on the Value of a Statistical Life," *Working Paper*, Department of Economics, Andrew Young School of Policy Studies, Georgia State University, September 2000; available as HTTP://DMSWEB.BADM.SC.EDU/GLENN/SCOTTONANDTAYLOR.PDF.

StataCorp, *Stata Statistical Software: Release 6.0* (College Station, TX: Stata Corporation, 1999).

United States Environmental Protection Agency, *The Benefits and Costs of the Clean Air Act, 1970 to 1990*. Office of Policy Analysis and Review, Document # A.97.20, October 1997; available as HTTP://DMSWEB.BADM.SC.EDU/GLENN/CAAREPORT_FULL.PDF.

Viscusi, W. Kip, "The Value of Risks to Life and Health," *Journal of Economic Literature*, XXXI, December 1993, 1912-1946.

Question and Answer Period for Session VI

Reed Johnson, of Research Triangle Institute, asked George van Houtven why he didn't consider it a problem if someone were to redefine the risk for cancer (because she thought a cure might be found) but not, for example, dying in an automobile accident.

George van Houtven answered that ultimately we are trying to get at people's perceptions. If they perceive that there is going to be a cure for cancer in the future, then they are going to discount it more, he said. If that is really their perception, we should not try to correct for that up front for policy purposes, because that is a valid reason for discounting.

Melonie Williams, of the EPA, commented that part of the problem is the way that we explain latency to people. She noted that the single-dip model that is used in order to avoid the complications of cumulative exposure effects means that we have to talk about the risk in terms of the exposure at the time of the exposure – so we are valuing exposure today to a carcinogen that is going to cause a cancer that's going to manifest X years from now. It is not scenario rejection, it is just that we can't control for it.

George van Houtven added that they can and do hope to control for it by asking debriefing questions, including whether respondents thought there was a good chance that there would be a cure for cancer.

Laura Taylor, of Georgia State University, commented that in her work she found that objective measures (of environmental quality) explain price better than subjective measures when these measures were on the same scale, and that people, when buying a house, did relatively well at identifying environmental quality levels, compared to the objective measure. She suggested that this raises the possibility that we are actually responding to an objective stimulus (e.g., to an average risk grade or some notion of risk), but that we don't verbalize it.

Glenn Harrison responded that we might expect that, because people will evolve heuristics over time, such that they may behave *as if* they have responded to the objective risks. To test that hypothesis, he said, you have to measure subjective risks. It is critical, he added, to distinguish between subjective risks and *hypothetical* subjective risks (that we often see in the literature). He suggested an alternative to calculating a social VSL by summing up the individual VSLs based on subjective risks – namely, that we plug in the objective risk in place of the subjective risk and recalculate what individuals' valuations would be if they got the risk estimate right. He referred to this as "a nice hybrid" – it respects the individuals' valuations, but says they just got one of the parameters (e.g., the risk) wrong. In order to be able to do that, he noted, we need to properly assess subjective risks.

Laura Taylor asked which is better, computing VSLs based on subjective measures or objective measures?

Glenn Harrison responded that, if you want to correct for the misperception of risk, you can use the VSL based on the objective risk. But, he added, you should be explicit about that. He added that his own preference is to use the subjective risk.

Tom Crocker, of the University of Wyoming, made the point that, if people behave according to their subjective risks (i.e., what they believe the risks to be), that will influence the odds that they will be subject to whatever consequence it is – i.e., that their subjective risks will actually influence the objective risks to them. In that case, subjective and objective risks aren't independent of each other.

Glenn Harrison agreed that there is an interdependence. He clarified that by objective risk he means the observed fatality data, recognizing the errors in those data, and by subjective risk he means the values he elicits from people, with incentives.

Reed Johnson suggested that people may choose not to invest in acquiring information about low probability risks because they do not care about them, so that perhaps what we characterize as mistakes they are making may in some sense be optimal.

Glenn Harrison agreed. He noted that this, however, runs afoul of one of the recommendations of the NOAA panel (with which he disagreed). The NOAA panel, he noted, recommends that we should make sure that the subjects in our CVM surveys are at least as intelligent, at least as well-informed, as the best-informed person. Harrison expressed the opinion that recommendation is elitist and violates the principle (in economics) of consumer voter sovereignty.

Kerry Smith, of North Carolina State University, noted that people will invest different amounts of time to find out the information necessary to assess the subjective probabilities depending on how important it is to them. If the incentive structure they are given (in a study) to behave correctly does not correspond to the incentives in the real world, then we are getting them to respond based on how the study rewards them – and that is not what is relevant to their actual behavior. He suggested, as an alternative, an approach taken by Evans and Viscusi (in the Review of Economics and Statistics), in which they looked at people's choices and tried to go backwards and infer what the probabilities had to have been in order for those choices to be made. With regard to the NOAA panel recommendation, he said that the panel was not making recommendations about how we should do stated preference research in general, but rather how we should do stated preference research when we are trying to compute compensation.

Glenn Harrison responded that the general failing of the NOAA report was to make prescriptions without providing any coherent or intelligent rationale. The very act of providing information, he noted, can influence the valuation, and we should be interested in both the ex ante and ex post evaluation. Uncertainty or lack of information, he said, does not mean we cannot value something; it simply means the valuation has more noise than when we have been given more information. We should not discount that initial valuation, he suggested, because that is what ought to be compensated.

Michael Hanemann, of the University of California, Berkeley, commented that what comes out of the three papers presented in this session is that the weak link in estimating VSL is the probability of risk. Everything else is okay. A major issue with subjective probability, he noted, is not just that it may differ from the objective risk (e.g., the risk is one in 10,000 but people think it's three in 10,000). There is the whole question of the metric with which people see probabilities. He offered the example that one may recognize that three in 10,000 is a larger number than one in 10,000, but may feel those are both small numbers, and so in a sense they are the same number. It is the idea in prospect theory, he said, that we need the mapping from objective to subjective. Part of that mapping, he noted, is getting the right units for subjective risk. Subjective risk, he suggested, may be more like a step function than continuous – that is, people may see things in risk classes (e.g., "a pretty big risk" or "a pretty small risk"). The problem, he said, is to fathom the units in which risks are perceived and communicate in those units.

Glenn Harrison agreed. His only concern, he said, is that the behaviorists have generated data that is worthless and that cannot be used to infer what that mapping is.

Ted Miller, of the Pacific Institute for Research and Evaluation, offered several caveats: (1) through some oddities in ICD-9 ordering of choice of death, "pneumonia" usually is actually either a heart disease, a cancer, or a hip fracture that through precedence rules gets coded as a respiratory death; (2) in1999, the precedence rules changed, resulting in a big drop in pneumonia deaths and an increase in hip fractures and heart disease; and (3) the ICD codes change over the years. He also suggested that violent deaths (e.g., murders) might be treated separately because of the dread factor. Finally, he asked Glenn Harrison for clarification on what deaths he was looking at.

Glenn Harrison said he wanted to look at all causes of death, but to break them out and look at them jointly.

Ted Miller asked how he does willingness to pay out of that – what are we paying for?

Glenn Harrison responded that it's just a hedonic wage equation – just as normally we would put fatal and non-fatal risk, he is simply decomposing fatal risk.

Ted Miller asked Glenn Harrison if he was expecting that people's occupation and their wage rate will explain all these things.

Glenn Harrison responded that there may be a weaker correlation between occupation and industry, and deaths from cancer, than there is from deaths on the job.

Ted Miller noted that this gets into problems. First, people sort themselves into jobs – a problem, he noted, that already exists in the wage risk studies. As an example, he said, there is data that suggest that people take jobs as bartenders and cooks if they like to drink on the job, and that drinking on the job is a risk of death from a lot of causes. So there is a correlation, he

said, because a person selects a place where the work rules let him behave in a certain way.

Glenn Harrison described this as the endogeneity of risk. He agreed that the more broad the death risks that one looks at, the more one has to worry about other factors – e.g., the smoking decision can affect the risk of lung cancer. He noted that the researcher has to endogenize the choice of risk, and he said that is very much something they are doing.

Ted Miller noted that another problem is that the data sets Glenn Harrison is using have near-term occupations, but not a history of occupations.

Glenn Harrison responded that they do give usual occupation during your life.

Ted Miller said that "usual occupation" is a subjective sort of thing. And there's substantial misclassification.

Glenn Harrison disagreed about the significance of the misclassification of occupation issue.