Toward a General Model of Demand for Health, Morbidity and Mortality States Part I: A Theoretical Analysis

Trudy Ann Cameron and J.R. DeShazo[1,2]

Abstract

We generalize the standard model of demand for a statistical life by more fully describing different health states and by recasting the model in an option price framework. By doing so, we eliminate several sources of upward bias in the VSL that result from excluding substitution possibilities. Our model permits distinct estimates of the individual's marginal willingness to pay to avoid a sick-year and a prematurely lost life-year. Using this approach, it will be possible to evaluate aggregate demand for programs that shift the distribution of illness profiles, where an illness profile describes sequential periods of health, morbidity, possible recovery, and premature mortality. Focusing on illness profiles will significantly enhance the evaluation of public life-saving policies.

[1]Senior authorship is not assigned. Lead authorship will rotate though this project paper series.

[2]Department of Economics, University of Oregon, and Department of Policy Studies, UCLA, respectively. We thank Vic Adamowicz, Richard Carson, Maureen Cropper, Baruch Fischhoff, Jim Hammitt, Alan Krupnick, and V. Kerry Smith for helpful reviews of our survey instrument and/or comments on our early results. Rick Li implemented our survey very capably. Ryan Bosworth and Graham Crawford have provided able assistance. This research was encouraged by Paul De Civita of the Economic Analysis and Evaluation Division of the Healthy Environments and Consumer Safety Branch (HECS) of Health Canada and has been generously supported by Health Canada (Contract H5431-010041/001/SS) and by the US Environmental Protection Agency (R829485), although it has not yet been formally reviewed by ether agency. Any remaining errors are our own.

1 Introduction

Constructing and estimating models of individual demand for future health states is one of the most challenging tasks economists face. Traditionally, scholars have focused on the individual's marginal rate of substitution between two just health states, life and death. In the standard model, the individual considers her utility in both states, weighted by the probability that each of these states will occur (Dreze 1967; Jones-Lee, 1974). One may then calculate the individual's willingness to trade off wealth for a reduced probability of death, holding utility constant. This simple model has served as the theoretical basis for hundreds of empirical studies designed to estimate the value of a statistical life, including those used to evaluate the appropriateness of life-saving public policies. (Viscusi, 1993; Aldy and Viscusi, 2002; Mrozek and Taylor, 2002).

In the context of the consumer's problem, we generalize the standard model in two ways. Subject to an income constraint, the consumer's problem is one of choosing quantities of costly risk-management programs that reduce the likelihood of the undesirable health states. As a first generalization, rather than assuming the individual faces only two homogeneous health states, our model assumes the individual faces a large set of multi-dimensional and time-denominated health states. Second, rather than assuming the individual makes a single lump-sum payment for a program in the current period, we recast the individual's demand for risk management programs in an option price framework. This accommodates programs with a future stream of certain costs and uncertain benefits.

Our more general model offers numerous advantages over the conventional specification. First, it eliminates several currently problematic sources of upward bias in estimates of the VSL. These upward biases are present when scholars omit relevant substitute states or dimensions of states (such as time) along which substitution may occur. Second, our model enables scholars to estimate the demand for programs that affect any feasible time sequence of states that involves periods of health, morbidity, or mortality. Third, because both the costs and health states are time-denominated, our model produces estimates of individual marginal willingness to pay to avoid a sick-year and prematurely lost life-year.

The first step in our model's theoretical development is to more completely describe the set of health states and the time-sequence in which the individual faces them. To begin with, we allow for a wide range of morbidity states, in addition to states of perfect health and sudden death. The reference state, from which future states are being evaluated, may be a state of morbidity rather than one of perfect health. The admission of morbidity states to the individual's choice set also allows for the possibility that there may be fates (states) worse than death for some individuals. Failing to recognize either of these cases, when they hold, is likely to bias upward estimates of VSL.

We also allow individuals to express their willingness to pay to avoid concatenations of health states, which we call "illness profiles." Each profile is defined as a probabilistic sequence of health states in the context of the individual's current age and remaining nominal lifespan. Each profile describes the future age of onset of an illness, sudden death if it occurs, and if it does not, the level and duration of pain and disability that follow. The remainder of the profile is described in terms of either full recovery or premature death after a number of years of pain and disability.

Describing future health states in terms of illness profiles is useful for several reasons. First, individuals actually think and talk about future health states in terms of illnesses (and their health-state profiles). Health profiles offer the appropriate lexicon and unit of analysis. Second, many risk management programs target specific illnesses. When choosing across programs, we allow individuals to substitute across illness profiles. Third, when choosing the quantity of a program to consume, it is quite likely that the individual is substituting among periods of health states within an illness profile. Our model explicitly allows for inter-temporal substitution among health states within an illness profile. We recover individual preferences over the latency periods associated with the onset of an illness, the duration of symptoms, the period of post-illness recovered status (if there is one), and the years lost to premature death. In contrast, most existing studies evaluate the individual willingness-to-pay for mortality risk reductions in the current year only. Fourth, our use of illness profiles captures the reality that mortality states are often positively correlated with morbidity states. VSL estimates from models that omit the associated morbidity will tend to be biased upward since the VSL from such a model conflates the value of avoided mortality and avoided morbidity.

Our second major theoretical contribution is the derivation of option prices for programs that reduce the risk of experiencing illness profiles. An option price framework accommodates choices involving the inter-temporal distribution of a future stream of certain costs and uncertain benefits (Graham, 1981). By explicitly discounting the future periods of costs and benefits, we are able to define the constant current-value payment that makes an individual just indifferent to an annual treatment program to reduce the risk of a future illness. When scaled to correspond to a 1.00 change in risk, the present discounted value of these payments is the multi-period optimization analog to a conventional VSL.

Our theoretical model, which is cast in terms of the indirect utility function, is designed to support the empirical phase of this research project. In this conjoint choice study, each illness profile is characterized as one of twelve illnesses which represent the major causes of death in society today. Each illness profile in the conjoint choice exercise is described in richer detail than will be carried through the devlopment of the following theoretical model.

Together, our theoretical framework and the forthcoming empirical analysis will provide two methodological contributions to the literature on valuing future health states. First, our option price formulas lead naturally to what we have labeled as the "value of a statistical illness" (VSI). Any given illness is associated with a distribution of specific profiles that a given population will experience (e.g., morbidity with full recovery, chronic morbidity with no premature death, morbidity followed by premature death, and sudden premature death). When a public policy reduces the incidence of an illness, it is altering the joint distribution of these illness profiles in the population. Our framework captures the value of avoiding all of the illness profiles for a particular illness, not just those associated with sudden death.

Second, by applying the option price framework to an illness profile we offer an important methodological advance: we estimate the marginal value of a life-year in various health states. Specifically, we estimate the marginal value of an additional life year spent in illness, spent in recovery, and lost to premature death. This innovation will contribute substantially to both the risk-valuation literature and, ultimately, to meeting the needs of policymakers. Policymakers often evaluate policies that postpone, but do not prevent, specific health outcomes. Until now they have had no estimates of the value of incrementally delaying the onset of an illness or death or of hastening recovery.

We organize this paper as follows. Section 2 present develops our option price model. Section 3 describes the broader theoretical construct of the value of a statistical illness and compares it to the value of the statistical life. In Section 4, we discuss two directions for future research before concluding in Section 5.

2 Theoretical Model: Indirect utility from an illness profile

Within the conventional VSL model, individuals may choose only two states–a "healthy" state or a "dead" state–in the current period. We expand upon these by considering four states: 1) a pre-illness healthy state, 2) illness state, 3) a post-illness recovered state and 4) a dead state.[3] We define each of these states as a time segment. Within each segment, the individual's health status is assumed (for now) to be relatively homogeneous.

2.1 Indirect utility from an illness profile

In Figures 1 and 2, we represent these four discrete health states. Let i index individuals and let t index time periods. To capture an illness profile, we use sets of dummy variables that collectively exhaust the period of time between the individual's present age and the end of his nominal life expectancy. The dummy variable $Pre\text{-}illness\_year_{it}$ take a value of 1 in years when the individual enjoys a healthy state. When the health state ends, the value of $Pre\text{-}illness\_year_{it}$ changes to 0 and remains there for the rest of the individual's expected lifespan. At the end of the healthy period the individual may die suddenly or become sick. Let the dummy variable $Illness\_year_{it}$ take on a value of 1 at this point and remain equal to 1 for the years during which the individual is ill. When he is not sick, it takes a value of zero. If the individual recovers from this illness, he may not recover to the exact state of health he experienced prior to the illness. The dummy variable labeled $Recover\_year_{it}$ takes on value of 1 in the years between the conclusion of the illness and the individual's expected time of death. Finally, we define $Lifeyear\_lost_{it}$ to distinguish the extent to which death is premature (that is, the time between death and what would otherwise have been

[3] Within our empirical model, the illness states are further differentiated into one of twelve specific illnesses, each of which can exhibit a wide variety of different symptom-treatment profiles that may last from zero to six years. In appropriate cases, the illness may also be chronic, lasting for more than six years.

the individual's nominal life expectancy).

For each health-state period, we assume initially that the indirect utility derived per unit of time from that particular health state is constant within that period. Time may be measured in years, months, or even a smaller units of time, depending on the degree of resolution needed. Next we define the future undiscounted indirect utilities per unit of time in each health state. Let these marginal utilities be denoted as δ_s for an episode of type s, where s in our model can be illness, recovered status, or a life-year lost to premature death. Let the undiscounted utility from each future year in a particular health state be defined *relative to no new illness*. In other words, we will normalize utility on the level of utility being experienced by the individual in their current health state.

We develop a simple linear model of the individual's future undiscounted indirect utility as a function of their health state in that future period. We abbreviate $\textit{Pre-illness\_year}_{it}$ to pre_{it}, $\textit{Illness\_year}_{it}$ to ill_{it}, $\textit{Recover\_year}_{it}$ to rcv_{it} and $\textit{Lifeyear\_lost}_{it}$ to lyl_{it} to allow more-compact notation.

$$V_{it} = \beta\left(Y_i\right) + \delta_1 ill_{it} + \delta_2 rcv_{it} + \delta_3 lyl_{it} + \eta_{it} \tag{1}$$

Let the undiscounted (dis)utility from each future year of illness be defined as δ_1, from each year of the post-illness recovered state be δ_2, and from each year of being prematurely dead be δ_3. The disutility of each of these states can be interpreted as being the same as the utility associated with avoiding them. The role of the dummy variables, ill_{it} , rcv_{it}, and lyl_{it}

will be simply to adjust the limits of the summations used for the present value of future continued good health, future intervals of illness, recovered time, and life-years lost. In this paper we assume that the individual uses the same discount rate, r, to discount both future money costs and the future disutility from either illness or premature mortality.[4]

With this set-up, we can develop a structural model of the ex ante option price that an individual will be willing to pay for a program that reduces his/her risk of a morbidity/mortality profile over the future. Define the present discounted value of indirect utility V_i^{jk} for the ith individual when $j = A$ if the program is chosen and $j = N$ if the program is not chosen. The superscript k will be S if the individual suffers the illness (or injury) and H if the individual does not suffer the illness.

The present value of indirect utility if the individual *does* choose the program and *does* suffer the illness takes the following form. All summations below will run from 0 to T_i, the remaining number of years in the individual's nominal life expectancy:

$$\begin{aligned} PDV(V_i^{AS}) &= \beta\left(Y_i\right)\sum\frac{\left(1-lyl_{it}^{A}\right)}{\left(1+r\right)^{t}}-\beta(c_i^{A})\sum\frac{\left(pre_{it}^{A}+rcv_{it}^{A}\right)}{\left(1+r\right)^{t}} \\ &\quad +\delta_1\sum\frac{ill_{it}^{A}}{\left(1+r\right)^{t}}+\delta_2\sum\frac{rcv_{it}^{A}}{\left(1+r\right)^{t}}+\delta_3\sum\frac{lyl_{it}^{A}}{\left(1+r\right)^{t}}+\varepsilon_i^{AS} \end{aligned} \tag{2}$$

[4] Empirically estimated discount rates for future money as opposed to future health states are suspected to differ to some extent. Discount rates also differ across individuals and across choice contexts, time horizons and sizes and types of outcomes at stake. No comprehensive empirical work has been undertaken that conclusively demonstrates the relationships between money and heath discount rates.

If we were to choose hyperbolic discounting for our specification, all of the discount factors in the expressions for present discounted value, below, would need to be changed from $1/(1+r)^t$ to $1/(1+t)^\lambda$. Other than this, the formulas will be the same.

What the individual assumes about their future income and program costs, if they choose the program, has implications for the formulas we develop in later sections. For their future income, our default assumption will be that individuals expect constant real annual income Y_i in each future year until the expected time of death if the individual gets the illness. The term $\left(1 - lyl_{it}^{A}\right)$ in equation (2) will be nonzero in those periods when the individual is still alive. While earned income is likely to suffer if the individual gets the illness, we assume that their annual income can be sustained through insurance coverage. For program costs, we assume that the annual costs of the risk-management program in question are incurred in the years leading up to the onset of the illness or injury, but are not paid while the individual is sick or injured.[5] If the individual recovers from the illness or injury, rather than dying from it, they will again participate in the risk-management program until their death. The term $pre_{it}^{A} + rcv_{it}^{A}$ in equation (2) will be non-zero only prior to the onset of the illness or during the recovered state.

The present value indirect utility, if the individual *does* choose the program but *does not* suffer the illness, involves no illness, recovery, or reduced lifespan. Thus, the expression for indirect utility takes the following form:

$$PDV(V_i^{AH}) = \beta\left(Y_i - c_i^A\right)\sum\frac{1}{(1+r)^t} + \varepsilon_i^{AH} \tag{3}$$

[5] While the individual is sick, the health testing program would provide no further information, and we assume that the major traffic accident is likely to result in the vehicle being "totaled" so that a new vehicle, with its safety features, would not be acquired until the individual has recovered from his or her injuries.

In this case, both income and the annual costs of program will continue until the end of the individual's nominal life expectancy. However, there are no benefits in the form of illness-years or lost life-years avoided.

Present value indirect utility, if the individual *does not* choose the program but *does* suffer the illness, is given by:

$$\begin{aligned} PDV(V_i^{NS}) &= \beta\left(Y_i\right)\sum\frac{\left(1-lyl_{it}^{A}\right)}{\left(1+r\right)^t} \\ &\quad +\delta_1\sum\frac{ill_{it}^{A}}{\left(1+r\right)^t}+\delta_2\sum\frac{rcv_{it}^{A}}{\left(1+r\right)^t}+\delta_3\sum\frac{lyl_{it}^{A}}{\left(1+r\right)^t}+\varepsilon_i^{NS} \end{aligned} \tag{4}$$

The individual's lifespan is potentially reduced, so future income continues only until the time of death, and the disutility of the illness, any recovery period, and any life-years lost will be relevant.

Present value indirect utility, if the individual *does not* choose the program and *does not* suffer the illness, is:

$$PDV(V_i^{NH})=\beta\left(Y_i\right)\sum\frac{1}{\left(1+r\right)^t}+\varepsilon_i^{NH} \tag{5}$$

Recall, the individual assumes that his current income level will be sustained until the end of his lifespan in the absence of premature mortality.

2.2 Expected indirect utility

In deriving the individual's option price for the program, given the ex ante uncertainty about future health states, we need to calculate *expected utilities.* In this case, the expectation is taken across the binary uncertain outcome of getting sick, S, or remaining healthy, H. The probability of illness or injury differs according to whether the respondent participates in the risk-reducing intervention program. Let the baseline probability of illness be Π_i^{NS} if the individual opts out of the program, and let the reduced probability be Π_i^{AS} if the individual opts in. The risk change due to program participation, $\Delta\Pi_i^{AS}$, is presumed to be negative.

Expected utility if the individual *buys* program A is:

$$E\left[V_i^A\right]_{S,H} = \Pi_i^{AS} \times PDV(V_i^{AS}) + \left(1 - \Pi_i^{AS}\right) \times PDV(V_i^{AH}) \tag{6}$$

$$= \Pi_i^{AS} \left[\begin{array}{l} \beta\left(Y_i\right) \sum \frac{\left(1-lyl_{it}^A\right)}{(1+r)^t} - \beta(c_i^A) \sum \frac{\left(pre_{it}^A + rcv_{it}^A\right)}{(1+r)^t} \\ +\delta_1 \sum \frac{ill_{it}^A}{(1+r)^t} + \delta_2 \sum \frac{rcv_{it}^A}{(1+r)^t} + \delta_3 \sum \frac{lyl_{it}^A}{(1+r)^t} + \varepsilon_i^{AS} \end{array} \right]$$
$$+ \left(1 - \Pi_i^{AS}\right) \left[\beta\left(Y_i - c_i^A\right) \sum \frac{1}{(1+r)^t} + \varepsilon_i^{AH}\right]$$

Expected utility if the program is *not* purchased (i.e. "no program", N), with the expectation taken over uncertainty about whether the individual will suffer the illness, is:

$$
\begin{aligned}
& E\left[V_i^N\right]_{S,H} = \Pi_i^{NS} \times PDV(V_i^{NS}) + \left(1 - \Pi_i^{NS}\right) \times PDV(V_i^{NH}) \qquad (7) \\
= \; & \Pi_i^{NS} \left[\begin{array}{l} \beta\left(Y_i\right) \sum \frac{1}{(1+r)^t} \\ \qquad + \delta_1 \sum \frac{ill_{it}^A}{(1+r)^t} + \delta_2 \sum \frac{rcv_{it}^A}{(1+r)^t} + \delta_3 \sum \frac{lyl_{it}^A}{(1+r)^t} + \varepsilon_i^{NS} \end{array} \right] \\
& + \left(1 - \Pi_i^{NS}\right) \left[\beta\left(Y_i\right) \sum \frac{1}{\left(1+r\right)^t} + \varepsilon_i^{NH}\right]
\end{aligned}
$$

By simplifying and collecting terms (See Appendix A), we can express the expected utility difference as:

$$
\begin{aligned}
& E\left[V_i^A\right] - E\left[V_i^N\right] \qquad (8) \\
= \; & \left[\beta\left(-c_i^A\right) \sum \frac{1}{\left(1+r\right)^t}\right] \\
& - \Pi_i^{AS} \left[\beta\left(Y_i\right) \sum \frac{lyl_{it}^A}{\left(1+r\right)^t}\right] - \Pi_i^{AS} \left[\beta(-c_i^A) \sum \frac{\left(1 - pre_{it}^A - rcv_{it}^A\right)}{\left(1+r\right)^t}\right] \\
& + \left(\Pi_i^{AS} - \Pi_i^{NS}\right) \left[\delta_1 \sum \frac{ill_{it}^A}{\left(1+r\right)^t} + \delta_2 \sum \frac{rcv_{it}^A}{\left(1+r\right)^t} + \delta_3 \sum \frac{lyl_{it}^A}{\left(1+r\right)^t}\right] + \varepsilon_i
\end{aligned}
$$

In this expression, the second line to the right of the equals sign is an artifact of our assumption about future income and the cost commitment implied by a choice of the program. The first term in this line corrects for the loss of income (meaning no consumption of other goods and services) after death. The second term in this line accommodates the presumption that program costs will not be paid if the individual is ill or dead.

Now define new variables based on the present value terms appearing in the formulas derived above:[6]

$$\begin{aligned} pdvc_i^A &= \sum \frac{1}{(1+r)^t} \\ pdvi_i^A &= \sum \frac{ill_{it}^A}{(1+r)^t} \\ pdvr_i^A &= \sum \frac{rcv_{it}^A}{(1+r)^t} \\ pdvl_i^A &= \sum \frac{lyl_{it}^A}{(1+r)^t} \end{aligned}$$

Equation (8) can then be written more compactly as:

$$\begin{aligned} & E\left[V_i^A\right] - E\left[V_i^N\right] \qquad (9) \\ = \; & \left[\beta\left(-c_i^A\right) pdvc_i^A\right] - \Pi_i^{AS}\left[\beta\left(Y_i\right) pdvl_i^A\right] - \Pi_i^{AS}\left[\beta(-c_i^A)\left(pdvi_i^A + pdvl_i^A\right)\right] \\ & + \left(\Delta\Pi_i^{AS}\right)\left[\delta_1 pdvi_i^A + \delta_2 \; pdvr_i^A + \delta_3 pdvl_i^A\right] + \varepsilon_i \end{aligned}$$

where $\Delta\Pi_i^{AS} = \Pi_i^{AS} - \Pi_i^{NS}$, which will be a negative number in our choice scenarios, since each involves a risk reduction.

[6] We refer to these compound explanatory variables using just the abbreviation for the main variable upon which each is based. Thus "pdvc" is used to denote $\left(-c_i^A\right)\left(pdvc_i^A - \Pi_i^{AS}\left(pdvi_i^A + pdvl_i^A\right)\right) - \left(Y_i\right)\Pi_i^{AS} pdvl_i^A$, "pdvi" to denote $\left(\Delta\Pi_i^{AS}\right) pdvi_i^A$, and so on.

2.3 Ex ante option prices

It will be convenient to be able to isolate the terms in c_i^A when we are calculating the values of statistical illnesses. We will also find it helpful to express the expected utility difference as linear-in-parameters, so equation (9) can be re-written as:

$$
\begin{aligned}
& E\left[V_i^A\right] - E\left[V_i^N\right] \qquad (10) \\
= \; & \beta \left[\; \left(-c_i^A\right) \left(pdvc_i^A - \Pi_i^{AS} \left(pdvi_i^A + pdvl_i^A\right)\right) \; - \; \left(Y_i\right) \Pi_i^{AS} pdvl_i^A \right] \\
& + \delta_1 \; \left(\Delta \Pi_i^{AS}\right) pdvi_i^A \\
& + \delta_2 \; \left(\Delta \Pi_i^{AS}\right) pdvr_i^A \\
& + \delta_3 \; \left(\Delta \Pi_i^{AS}\right) pdvl_i^A + \varepsilon_i
\end{aligned}
$$

The discrete choice among program alternatives can be modeled as depending upon the marginal utility of income, β, and the marginal utilities of each year of avoided degraded health status or premature mortality: δ_1, δ_2, and δ_3.[7] The variables that must be constructed in order to estimate these three key parameters are revealed in equation (10). We see from equation (10) that the difference in expected present value indirect utilities associated with choosing a risk-reduction program is a function of the illness profile as captured by the $pdvi_i^A$, $pdvr_i^A$,and $pdvl_i^A$ terms.

With respect to our empirical application, equation (10) is the basis for estimation of the

[7] In future analysis, we plan to let the choice depend upon individual specific discount rates.

random utility choice model that explains individuals' choices among the three alternatives presented in each choice scenario: Program A, Program B, or Neither Program. There is an analogous difference in expected utilities between Program B and the Neither Program choice. All choices posed to respondents were three-way choices, so the models will be estimated using McFadden's conditional logit estimator (or appropriate modifications of this model).

The option price for the program that accomplishes this decrease in illness probabilities is the common certain payment, regardless of which way the uncertainty about contracting the illness is resolved, that makes the individual just indifferent between paying for the program and enjoying the risk reduction, or not paying for the program and not enjoying the risk reduction. This payment, c_i^{A*}, will make $E\left[V_i^A\right] - E[V_i^N\right] = 0$. Setting equation (10) equal to zero and solving yields:

$$c_i^{A*} = \frac{\left[\begin{array}{l} \beta\ \left[-\left(Y_i\right)\Pi_i^{AS} pdvl_i^A\right] \\ \qquad +\delta_1\ \left(\Delta\Pi_i^{AS}\right) pdvi_i^A + \delta_2\ \left(\Delta\Pi_i^{AS}\right) pdvr_i^A + \delta_3\ \left(\Delta\Pi_i^{AS}\right) pdvl_i^A + \varepsilon_i \end{array}\right]}{\beta\ \left[pdvc_i^A - \Pi_i^{AS}\left(pdvi_i^A + pdvl_i^A\right)\right]} \tag{11}$$

where $\Delta\Pi_i^{AS} = \Pi_i^{AS} - \Pi_i^{NS}$ is the size of the risk reduction to be derived from participating in the program. The amount of money c_i^{A*} is the maximum constant annual payment that the individual will be willing to make, regardless of whether he suffers the illness, in order to purchase the program that reduces his probability of suffering the illness from Π_i^{NS} to Π_i^{AS}.

While the payment c_i^{A*} is the maximum annual payment the individual is willing to make, these payments are necessary for the rest of the individual's life, so the present value of these payments must be calculated. In this context, however, there is some uncertainty over just what will constitute "the rest of the individual's life," since this may differ according to whether the individual suffers the illness or not. We will use the expected present value of this time profile of costs, with the expectation taken over whether or not the individual suffers the illness when they are participating in the program.

$$\begin{aligned}
& E\left[PV(c_i^{A*})\right] \qquad (12) \\
= \; & \left(1-\Pi_i^{AS}\right)\left(c_i^{A*}\right)\sum\frac{1}{(1+r)^t}+\left(\Pi_i^{AS}\right)\left(c_i^{A*}\right)\sum\frac{\left(pre_{it}^A+rcv_{it}^A\right)}{(1+r)^t} \\
= \; & \left(c_i^{A*}\right)\sum\frac{1}{(1+r)^t}-\left(\Pi_i^{AS}\right)\left(c_i^{A*}\right)\sum\frac{1}{(1+r)^t}+\left(\Pi_i^{AS}\right)\left(c_i^{A*}\right)\sum\frac{\left(pre_{it}^A+rcv_{it}^A\right)}{(1+r)^t} \\
= \; & \left(c_i^{A*}\right)\left[\sum\frac{1}{(1+r)^t}-\left(\Pi_i^{AS}\right)\sum\frac{\left(1-pre_{it}^A-rcv_{it}^A\right)}{(1+r)^t}\right] \\
= \; & \left(c_i^{A*}\right)\left[pdvc_i^A-\Pi_i^{AS}\left(pdvi_i^A+pdvl_i^A\right)\right]
\end{aligned}$$

The expected present value of the lifetime stream of payments, given that the individual participates in the program, is therefore simply:

$$\begin{aligned}
& E\left[PV(c_i^{A*})\right] \qquad (13) \\
= \; & [\beta]^{-1}\left[\begin{array}{l} \beta\ \left[-\left(Y_i\right)\Pi_i^{AS}pdvl_i^A\right] \\ \qquad +\delta_1\ \left(\Delta\Pi_i^{AS}\right)pdvi_i^A+\delta_2\ \left(\Delta\Pi_i^{AS}\right)pdvr_i^A+\delta_3\ \left(\Delta\Pi_i^{AS}\right)pdvl_i^A+\varepsilon_i \end{array}\right]
\end{aligned}$$

2.4 Value of a statistical illness (VSI)

The expected present discounted value in equation (13) pertains to the maximum annual willingness to pay for a small risk reduction, $\Delta\Pi_i^{AS}$. To convert this to the value for a statistical illness (VSI), we divide by the absolute size of the risk reduction in order to scale this present value to the present discounted willingness to pay that would correspond to a 1.00 change in the risk.

$$
\begin{aligned}
&\frac{E\left[PV(c_i^{A*})\right]}{\left|\Delta\Pi_i^{AS}\right|} && (14)\\
= \;& \left[\beta\left|\Delta\Pi_i^{AS}\right|\right]^{-1}\left[\begin{array}{l}\beta\ \left[-\left(Y_i\right)\Pi_i^{AS} pdvl_i^A\right] \\ \quad +\delta_1\ \left(\Delta\Pi_i^{AS}\right) pdvi_i^A+\delta_2\ \left(\Delta\Pi_i^{AS}\right) pdvr_i^A+\delta_3\ \left(\Delta\Pi_i^{AS}\right) pdvl_i^A+\varepsilon_i\end{array}\right] \\
= \;& \left(\frac{-\left(Y_i\right)\Pi_i^{AS} pdvl_i^A}{\left|\Delta\Pi_i^{AS}\right|}\right)+\left(\frac{\delta_1\left(\Delta\Pi_i^{AS}\right)}{\beta\left|\Delta\Pi_i^{AS}\right|}\right) pdvi_i^A+\left(\frac{\delta_2\left(\Delta\Pi_i^{AS}\right)}{\beta\left|\Delta\Pi_i^{AS}\right|}\right) pdvr_i^A \\
& \qquad +\left(\frac{\delta_3\left(\Delta\Pi_i^{AS}\right)}{\beta\left|\Delta\Pi_i^{AS}\right|}\right) pdvl_i^A+\left(\frac{\varepsilon_i}{\beta\left|\Delta\Pi_i^{AS}\right|}\right)
\end{aligned}
$$

In our study, all probability changes $\Delta\Pi_i^{AS}$ are negative, while the absolute magnitude of these changes will be positive. Multiplication by $\Delta\Pi_i^{AS}/\left|\Delta\Pi_i^{AS}\right|$ will amount to multiplying by -1, which will change the effective sign on each of the terms involving this ratio. The effective formula for the value of a statistical illness will be:

$$
\begin{aligned}
VSI_i^A &= \frac{E\left[PV(c_i^{A*})\right]}{\left|\Delta\Pi_i^{AS}\right|} \qquad (15) \\
&= \left(\frac{-\delta_1}{\beta}\right) pdvi_i^A + \left(\frac{-\delta_2}{\beta}\right) pdvr_i^A + \left(\frac{-\delta_3}{\beta} - Q_i\right) pdvl_i^A + \left(\frac{\varepsilon_i}{\beta\left|\Delta\Pi_i^{AS}\right|}\right)
\end{aligned}
$$

where $Q_i = (Y_i)\,\Pi_i^{AS}/\left|\Delta\Pi_i^{AS}\right|$. Across the distribution of the logistic error term, ε_i, the expectation is zero, so the expected value of a statistical illness depends only on the systematic portion of equation (14). The VSI will depend upon the different marginal utilities of avoided periods of illness, recovered status, and premature death. It will also depend upon the time profiles for each of these states as embedded in the terms $pdvi_i^A$, $pdvr_i^A$, and $pdvl_i^A$, and (implicit in this model) upon the individual's own discount rate.[8]

Q_i consists only of data and must be strictly positive. Any non-zero Q_i will shrink the predicted point estimate of the VSI. The shrinkage will be greater (i.) as income is larger, (ii.) as more life-years are lost, (iii.) as the individual is older, so that life-years lost come sooner in time, (iv.) the larger the remaining risk, and (v.) the smaller the absolute risk reduction.[9]

[8] Subsequent work will preserve individual discount rates as systematically varying parameters, to be estimated with reference to the individual's responses to a hypothetical lottery question. Here, discount rates are presumed to be exogenous and constant across individuals. Our empirical work explores the consequences of using different discount rate assumptions.

[9] Nothing in this specification precludes negative point estimates of the VSI. It is possible to estimate a positive value for the marginal utility of income, β, and negative values for the marginal utilities of illness-years, recovered-years, and lost life-years (the δs). This would guarantee a positive fitted value for VSI if $Q_i = 0$ as in the naive-costs case. If we assume a sophisticated cost interpretation, however, for Q_i sufficiently large, the point estimate of VSI can be negative, even with homogeneous preferences.

The key undiscounted marginal utility parameters are not presently constrained to be strictly positive

In expectation, the fitted value of a statistical illness can potentially vary systematically across types of illnesses according to the labels assigned to the illnesses, the symptoms and treatment associated with them, the individual's characteristics (such as age and gender), perceptions of risks associated with the type of illness, and prior experience with that illness. This heterogeneity can be accommodated by making the indirect utility parameters δ_1, δ_2, and δ_3 depend upon individual characteristics or attributes of each illness.

The error term ε in equation (15) is assumed to be identically distributed across observations in a manner appropriate for conditional logit estimation. Given the transformation needed to solve for the VSI, however, the error term in the VSI formula will be heteroscedastic, with smaller error variances corresponding to cases with larger absolute risk reductions, $\left|\Delta\Pi_i^{AS}\right|$.

In our future empirical application, the addition of illness labels and a symptom-treatment profile (within the illness state) will convey to the respondent some information about what health consequences might ensue from each illness we describe. These illness characteristics can be expected to shift the value of δ_1, the marginal (dis)utility of a sick-year. The marginal utility of each period of recovered health status, δ_2, could be allowed also to vary by type of illness as well, since the illness labels may connote the degree of "health" that nominal

(for income) and strictly negative (for episodes of undesirable health states). This is especially a concern when these marginal utilities are permitted to vary systematically with of the attributes of the illness profile and/or the characteristics of the individual in question. The marginal utility of income, the scalar parameter β in our simplest models, bears a point estimate that is robustly positive, but positive values for one or both of the systematically varying parameters capturing the marginal utility of an illness-year (δ_1) or a lost life-year (δ_3) can push an individual fitted value of the VSI for a particular morbidity/mortality profile into the negative range.

recovery from that illness actually implies. Finally, the marginal utility of a lost life-year may depend upon the health state prior to death.

2.5 VSIs versus Conventional VSLs

The existing literature, especially the hedonic wage-risk literature, focuses on society's willingness to pay for incremental reductions in the chance of a sudden accidental death in the current period. In the framework of our illness profiles, such an event would be captured by zero years of morbidity and sudden death in the current year, with the remainder of the individual's nominal life expectancy experienced as lost life-years. In this case, a point estimate of the VSI would be given by equation (15), for a specification where a dummy variable for *die_suddenly* is allowed to shift the marginal utility of a lost life-year. Let δ_{30} be the baseline marginal utility of a lost life-year and let δ_{31} be the estimated coefficient on the dummy variable activated when death is sudden. Since the terms in $pdvi_i^A$ and $pdvr_i^A$ will be zero, our analog to the conventional VSL formula will be:

$$\begin{aligned} E[VSL] &= \frac{E\left[PV(c_i^{A*})\right]}{|\Delta\Pi_i^{AS}|} = \left(\frac{-(\delta_{30}+\delta_{31})}{\beta} - Q_i\right) pdvl_i^A, \text{ where} \qquad (16) \\ pdvl_i^A &= \sum \frac{lyl_{it}^A}{(1+r)^t} \text{ and } Q_i = \frac{(Y_i)\,\Pi_i^{AS}}{|\Delta\Pi_i^{AS}|} \end{aligned}$$

The summation in the formula for $pdvl_i^A$ is from the present until the individual's nominal life expectancy. This interval depends upon the individual's current age, so even in a model with homogeneous preferences, the VSI will vary with age. The VSI also depends upon the individual's income, on the absolute magnitude of the risk reduction, and on the remaining health risk with the program in place. Of course, the individual's discount rate will also matter.

2.6 Calculating Policy-Relevant VSIs

To be clear on what is needed to construct VSIs using our present results, we offer the following checklist:

1. For the illness in question: An approximate joint distribution for the possible ages of onset, possible durations of moderate and severe pain and disability, possible durations of hospitalization, possibilities of minor or major surgery, and possible reductions in lifespans, and possible outcomes (recovery, sudden death, morbidity less than six years, chronic morbidity more than six years). In practice, this joint distribution will be constructed using expert judgment and its validity will in part determine the validity of the eventual VSI estimates our model will produce.

2. For the population affected by this health threat: An approximate joint distribution of at least age, gender, and income level (and possibly other variables if richer models prove to provide robust results with our data). The distribution

of these characteristics may be based on expert judgment combined with exposure and epidemiological data. Again, the validity of the assumptions underlying this approximate joint distribution will in part determine the validity of the resulting VSI estimates.

3. Make a large number of random draws (say, 10,000 or even 1,000,000) from the joint distribution of illness profile characteristics (some of these characteristics may be judged to be independently distributed, others will be correlated.)

4. If illness profiles are independent of the age, gender and other characteristics of the affected population, make an equal number of random draws from the joint distribution of characteristics of the affected population. If the joint distribution of illness profiles differs by gender or by age category or income level, such joint illness distributions might preferably be specified for each category within the population, stratified by age, gender, and possibly income level.

5. Combine these illness attributes and individual characteristics with our derived formulas for the value of a statistical illness.

6. Generate a marginal distribution for the range of implied VSIs. The mean of this distribution can be interpreted as our model's prediction about the population average of VSIs for this type of health threat.

The overall Value of a Statistical Illness, estimated in this fashion and calculated for a given policy by simulation methods, will allow the researcher to more fully capture the policy choice context for the risk in question.

3 Two Directions for Future Research

3.1 The death event

Figures 1 and 2 display morbidity/mortality profiles that exhibit no particular recognition of the time of death. It is possible that the event of death bears greater significance for respondents than simply a transition from the status of being either healthy, sick, or recovered to being dead. To introduce the possibility that there is some additional disutility associated with the future event of dying, we can introduce another indicator variable, called $Death\_year_{it}$, which takes on a value of 1 in the year of death (also the first year of the $Lifeyear\_lost_{it}$ interval). Figure 3 displays a morbidity/mortality profile that includes explicit recognition of the event of death at a particular point in the future. The present value of the single future year of when death occurs would bear the coefficient δ_d in our specification, so that the overall effect of death in year t would be the increment in present-value utility that comes from death, as well as the usual increment to present-value utility that is associated with a lost life-year.

$$\delta_d \frac{1}{(1+r)^t} + \delta_3 \frac{1}{(1+r)^t} = (\delta_d + \delta_3)\frac{1}{(1+r)^t} \tag{17}$$

For years subsequent to the year of death, the present value of utility would reflect only the disutility of life-years lost:

$$\delta_3 \frac{1}{(1+r)^t} \tag{18}$$

If the event of death confers greater disutility than simply the status of being dead, then we would expect the parameter δ_d to be nonzero and the sum $(\delta_d + \delta_3)$ to convey greater disutility than just δ_3, the disutility associated with another year of being prematurely dead in each subsequent year of what might otherwise have been the individual's life expectancy.

Incorporating the event of death into our model involves modifying many of the main equations presented above. Assume that program A is designed to reduce the probability of death occuring in year t^A, whereas if the individual does not get sick or injured, they will expect to die in year t^e. Equation (2) can be modified to read:

$$\begin{aligned} V_i^{AS} &= \beta(Y_i) \sum \frac{\left(1 - lyl_{it}^A\right)}{(1+r)^t} - \beta(c_i^A) \sum \frac{\left(pre_{it}^A + rcv_{it}^A\right)}{(1+r)^t} \\ &\quad + \delta_1 \sum \frac{ill_{it}^A}{(1+r)^t} + \delta_2 \sum \frac{rcv_{it}^A}{(1+r)^t} + \delta_3 \sum \frac{lyl_{it}^A}{(1+r)^t} + \delta_d \frac{1}{(1+r)^{tA}} + \varepsilon_i^{AS} \end{aligned} \tag{19}$$

If the individual chooses the program, but gets sick anyway, his utility level will be adversely affected by δ_d in the year t^A when he is likely to die if he suffers from the illness. In contrast if th individual chooses the program but does not suffer the illness, his utility will be modified to the following if the event of death is distinguished as having a distinct effect on utility:

$$V_i^{AH} = \beta \left(Y_i - c_i^A\right) \sum \frac{1}{(1+r)^t} + \delta_d \frac{1}{(1+r)^{t^e}} + \varepsilon_i^{AH} \tag{20}$$

where the event of death is now postponed until year t^e which is the individual's nominal life expectancy.

If the individual does not choose the program, he may suffer the illness and die earlier, which will make his utility

$$\begin{aligned} V_i^{NS} &= \beta \left(Y_i\right) \sum \frac{\left(1 - lyl_{it}^A\right)}{(1+r)^t} \\ &\quad + \delta_1 \sum \frac{ill_{it}^A}{(1+r)^t} + \delta_2 \sum \frac{rcv_{it}^A}{(1+r)^t} + \delta_3 \sum \frac{lyl_{it}^A}{(1+r)^t} + \delta_d \frac{1}{(1+r)^{t^A}} + \varepsilon_i^{NS} \end{aligned} \tag{21}$$

If the individual does not choose the program and managed to avoid suffering the illness or injury, his utility will be

$$V_i^{NH} = \beta \left(Y_i\right) \sum \frac{1}{(1+r)^t} + \delta_d \frac{1}{(1+r)^{t^e}} + \varepsilon_i^{NH} \tag{22}$$

With these amendments, the expression for the expected utlity difference between program and no program will include the following additional terms:

$$\begin{aligned}
&\Pi_i^{AS}\left[\delta_d \frac{1}{(1+r)^{t^A}}\right]+\left(1-\Pi_i^{AS}\right)\left[\delta_d \frac{1}{(1+r)^{t^e}}\right] \\
&-\Pi_i^{NS}\left[\delta_d \frac{1}{(1+r)^{t^A}}\right]-\left(1-\Pi_i^{NS}\right)\left[\delta_d \frac{1}{(1+r)^{t^e}}\right] \\
=\; &\left(\Pi_i^{AS}-\Pi_i^{NS}\right)\left[\delta_d \frac{1}{(1+r)^{t^A}}\right]-\left(\Pi_i^{AS}-\Pi_i^{NS}\right)\left[\delta_d \frac{1}{(1+r)^{t^e}}\right]
\end{aligned}$$

Working through similar algebra, the eventual estimating specification can be expressed comparably if we define an additional simplification:

$$pdvd_i^A=\left(\frac{1}{(1+r)^{t^A}}-\frac{1}{(1+r)^{t^e}}\right) \tag{23}$$

Then we can write the modified expected indirect utility difference as

$$\begin{aligned}
&E\left[V_i^A\right]-E\left[V_i^N\right] \qquad (24)\\
=\; &\beta\;\left[\left(-c_i^A\right)\left(pdvc_i^A-\Pi_i^{AS}\left(pdvi_i^A+pdvl_i^A\right)\right)-\left(Y_i\right)\Pi_i^{AS}pdvl_i^A\right] \\
&+\delta_1\;\left(\Pi_i^{AS}-\Pi_i^{NS}\right)pdvi_i^A \\
&+\delta_2\;\left(\Pi_i^{AS}-\Pi_i^{NS}\right)pdvr_i^A \\
&+\delta_d\;\left(\Pi_i^{AS}-\Pi_i^{NS}\right)pdvd_i^A \\
&+\delta_3\;\left(\Pi_i^{AS}-\Pi_i^{NS}\right)pdvl_i^A+\varepsilon_i
\end{aligned}$$

In subsequent work, we will explore the potential for there to be an added disutility from the event of death.

3.2 Age-at-Illness rather than age now

Our basic specification assumes that the undiscounted (dis)utility of a year of illness or injury is a constant (in the homogeneous specification) or depends upon a number of characteristics of the individual *at the time they are asked to make program choices.* One important generalization to explore is the possibility that the undiscounted disutility of a future year of illness or injury depends upon the age of the individual *at the time they are experiencing that year of illness or injury.* In this case, equation (1) can be rewritten as a function of the respondent's age at time t, age_t, rather than their current age, which we will use elsewhere in our model with heterogeneous preferences and will denote by age_0:

$$V_{it} = \beta\left(Y_i\right) + (\delta_{10} + \delta_{11}age_t)ill_{it}^A + (\delta_{20} + \delta_{21}age_t)rcv_{it}^A + (\delta_{30} + \delta_{31}age_t)lyl_{it}^A + \eta_{it} \tag{25}$$

As an example of the implications of the specification in equation (25) for the estimation process, equation (2) now becomes:

$$\begin{aligned} V_i^{AS} &= \beta\left(Y_i\right)\sum\frac{\left(1-lyl_{it}^{A}\right)}{\left(1+r\right)^t}-\beta(c_i^A)\sum\frac{\left(pre_{it}^{A}+rcv_{it}^{A}\right)}{\left(1+r\right)^t} \\ &+\delta_{10}\sum\frac{ill_{it}^{A}}{\left(1+r\right)^t}+\delta_{11}\sum\frac{(age_t)(ill_{it}^{A})}{\left(1+r\right)^t} \\ &+\delta_{20}\sum\frac{rcv_{it}^{A}}{\left(1+r\right)^t}+\delta_{21}\sum\frac{(age_t)(rcv_{it}^{A})}{\left(1+r\right)^t} \\ &+\delta_{30}\sum\frac{lyl_{it}^{A}}{\left(1+r\right)^t}+\delta_{31}\sum\frac{(age_t)(lyl_{it}^{A})}{\left(1+r\right)^t}+\varepsilon_i^{AS} \end{aligned} \tag{26}$$

We will thus define some new variable acronyms:

$$\begin{aligned} agepdvi_i^A &= \sum\frac{(age_t)(ill_{it}^{A})}{\left(1+r\right)^t} \\ agepdvr_i^A &= \sum\frac{(age_t)(rcv_{it}^{A})}{\left(1+r\right)^t} \\ agepdvl_i^A &= \sum\frac{(age_t)(lyl_{it}^{A})}{\left(1+r\right)^t} \end{aligned} \tag{27}$$

This implies that the eventual estimating specification will become a generalization of equation (10):

$$
\begin{aligned}
& E\left[V_i^A\right] - E\left[V_i^N\right] \qquad (28) \\
= \; & \beta \left[\;\; \left(-c_i^A\right)\left(pdvc_i^A - \Pi_i^{AS}\left(pdvi_i^A + pdvl_i^A\right)\right) \quad - \quad \left(Y_i\right)\Pi_i^{AS} pdvl_i^A \right] \\
& + \delta_{10} \; \left(\Delta\Pi_i^{AS}\right) pdvi_i^A + \delta_{11} \; \left(\Delta\Pi_i^{AS}\right) agepdvi_i^A \\
& + \delta_{20} \; \left(\Delta\Pi_i^{AS}\right) pdvr_i^A + \delta_{21} \; \left(\Delta\Pi_i^{AS}\right) agepdvr_i^A \\
& + \delta_{30} \; \left(\Delta\Pi_i^{AS}\right) pdvl_i^A + \delta_{31} \; \left(\Delta\Pi_i^{AS}\right) agepdvl_i^A + \varepsilon_i
\end{aligned}
$$

If the marginal utility of a health state is not a linear function of age, but (perhaps) a quadratic function or some other simple non-linear relationship, we might pursue models with not only the contemporaneous age as an interaction term with each health state dummy, but also the square of this variable. Or perhaps we will find that the logarithm of age has greater predictive power, or that age captured as several categories will work best.

4 Conclusions

Policy analysis with respect to risk-management programs requires detailed information about consumer demand for these programs. We have set out to build a formal utility-theoretic model that captures the relevant considerations in private ex ante consumer choices about incurring ongoing expenditures to reduce risks to life and health. Most past studies have focused on current-period costs and current-period benefits. In contrast, our model recognizes the future time profiles of illnesses and injuries for which individuals may choose to act to reduce their risks. Intertemporal consumer optimization requires explicit treatment of the interaction between disease latencies and individual discount rates. Our model permits us to derive option prices for programs that reduce well-defined types of risks. Option prices are the appropriate theoretical construct for decision-making under uncertainty, where the uncertainty in this case concerns whether the individual will actually suffer the illness or injury that the proposed risk reduction measure addresses.

We show that our option price formulas lead naturally to what we have labeled as the "value of a statistical illness" (VSI). The VSI is the present discounted value of the stream of maximum annual payments that the individual would be willing to pay for the specified (typically small) risk reduction, scaled up proportionately to correspond to a risk reduction of 100%. This construct is analogous to the more familiar, but more-limited, concept of the value of a statistical life (VSL). A VSL is typically constructed by looking simply at the static single-period willingness to pay for a specified risk reduction, and scaling this willingness to pay up to a 100% risk reduction. However, static VSL estimates do not typically vary with important morbidity/mortality attributes such as latency, time profiles of illness, symptoms and treatments, outcomes, or life-years lost.

5 Bibliography

Dreze, J. (1962) ""L'Utilite Sociale d'une Vie Humaine"." *Revue Francaise de Recherche Operationalle* 6: 93 - 118.

Jones-Lee, M. (1974) "The Value of Changes in the Probability of Death or Injury." *Journal of Political Economy* 82(4): 835 - 849.

Graham, D.A. (1981) "Cost-benefit Analysis Under Uncertainty." *American Economic Review* 71 (4): 715-725 1981

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

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

6 Appendix A

The difference in expected present discounted utility between the program and no-program cases will be:

$$
\begin{aligned}
& E\left[V_i^A\right] - E\left[V_i^N\right] \qquad (29) \\
= \ & \Pi_i^{AS} \left[\begin{array}{l} \beta\left(Y_i\right) \sum \frac{\left(1 - lyl_{it}^A\right)}{(1+r)^t} - \beta(c_i^A) \sum \frac{\left(pre_{it}^A + rcv_{it}^A\right)}{(1+r)^t} \\ \quad + \delta_1 \sum \frac{ill_{it}^A}{(1+r)^t} + \delta_2 \sum \frac{rcv_{it}^A}{(1+r)^t} + \delta_3 \sum \frac{lyl_{it}^A}{(1+r)^t} + \varepsilon_i^{AS} \end{array} \right] \\
& + \left(1 - \Pi_i^{AS}\right) \left[\beta\left(Y_i - c_i^A\right) \sum \frac{1}{(1+r)^t} + \varepsilon_i^{AH}\right] \\
& - \Pi_i^{NS} \left[\begin{array}{l} \beta\left(Y_i\right) \sum \frac{1}{(1+r)^t} \\ \qquad + \delta_1 \sum \frac{ill_{it}^A}{(1+r)^t} + \delta_2 \sum \frac{rcv_{it}^A}{(1+r)^t} + \delta_3 \sum \frac{lyl_{it}^A}{(1+r)^t} + \varepsilon_i^{NS} \end{array} \right] \\
& - \left(1 - \Pi_i^{NS}\right) \left[\beta\left(Y_i\right) \sum \frac{1}{(1+r)^t} + \varepsilon_i^{NH}\right]
\end{aligned}
$$

Distributing terms, this expected utility difference can be written as:

$$
\begin{aligned}
& E\left[V_i^A\right]-E\left[V_i^N\right] \qquad (30)\\
= \; & \Pi_i^{AS}\left[\beta\left(Y_i\right)\sum\frac{\left(1-lyl_{it}^A\right)}{(1+r)^t}-\beta(c_i^A)\sum\frac{\left(pre_{it}^A+rcv_{it}^A\right)}{(1+r)^t}\right] \\
& +\Pi_i^{AS}\left[\delta_1\sum\frac{ill_{it}^A}{(1+r)^t}+\delta_2\sum\frac{rcv_{it}^A}{(1+r)^t}+\delta_3\sum\frac{lyl_{it}^A}{(1+r)^t}\right]+\Pi_i^{AS}\left[\varepsilon_i^{AS}\right] \\
& +\left[\beta\left(Y_i-c_i^A\right)\sum\frac{1}{(1+r)^t}\right]-\Pi_i^{AS}\left[\beta\left(Y_i-c_i^A\right)\sum\frac{1}{(1+r)^t}\right] \\
& \qquad +\left(1-\Pi_i^{AS}\right)\left[\varepsilon_i^{AH}\right] \\
& -\Pi_i^{NS}\left[\beta\left(Y_i\right)\sum\frac{1}{(1+r)^t}\right] \\
& -\Pi_i^{NS}\left[\delta_1\sum\frac{ill_{it}^A}{(1+r)^t}+\delta_2\sum\frac{rcv_{it}^A}{(1+r)^t}+\delta_3\sum\frac{lyl_{it}^A}{(1+r)^t}\right]-\Pi_i^{NS}\left[\varepsilon_i^{NS}\right] \\
& -\left[\beta\left(Y_i\right)\sum\frac{1}{(1+r)^t}\right]+\Pi_i^{NS}\left[\beta\left(Y_i\right)\sum\frac{1}{(1+r)^t}\right]-\left(1-\Pi_i^{NS}\right)\left[\varepsilon_i^{NH}\right]
\end{aligned}
$$

In the process of simplifying this expression, there are four components involving error terms. We will define the compound error term as ε. If the error terms ε_i^{Nk} are independent and identically distributed according to an extreme value distribution, and if the ε_i^{Nk} are similarly independent and identically distributed extreme value, then the resulting error term can be assumed to be logistic, so that a logit model is appropriate.

$$
\varepsilon_i=\Pi_i^{AS}\left[\varepsilon_i^{AS}\right]+\left(1-\Pi_i^{AS}\right)\left[\varepsilon_i^{AH}\right]-\Pi_i^{NS}\left[\varepsilon_i^{NS}\right]-\left(1-\Pi_i^{NS}\right)\left[\varepsilon_i^{NH}\right] \qquad (31)
$$

In equation (30), the two terms involving Π_i^{NS} and β multiplying the present value of gross income will cancel. Two other component-wise simplifications include:

$$\left[\beta\left(Y_i - c_i^A\right)\sum\frac{1}{(1+r)^t}\right] - \left[\beta\left(Y_i\right)\sum\frac{1}{(1+r)^t}\right] = \left[\beta\left(-c_i^A\right)\sum\frac{1}{(1+r)^t}\right] \tag{32}$$

and

$$\begin{aligned}
& \Pi_i^{AS}\left[\beta\left(Y_i\right)\sum\frac{\left(1-lyl_{it}^A\right)}{(1+r)^t} - \beta(c_i^A)\sum\frac{\left(pre_{it}^A + rcv_{it}^A\right)}{(1+r)^t}\right] \\
& -\Pi_i^{AS}\left[\beta\left(Y_i - c_i^A\right)\sum\frac{1}{(1+r)^t}\right] \\
= \; & \Pi_i^{AS}\left[\beta\left(Y_i\right)\sum\frac{\left(1-lyl_{it}^A\right)}{(1+r)^t}\right] + \Pi_i^{AS}\left[\beta(-c_i^A)\sum\frac{\left(pre_{it}^A + rcv_{it}^A\right)}{(1+r)^t}\right] \\
& -\Pi_i^{AS}\left[\beta\left(Y_i\right)\sum\frac{1}{(1+r)^t}\right] - \Pi_i^{AS}\left[\beta\left(-c_i^A\right)\sum\frac{1}{(1+r)^t}\right] \\
= \; & -\Pi_i^{AS}\left[\beta\left(Y_i\right)\sum\frac{lyl_{it}^A}{(1+r)^t}\right] - \Pi_i^{AS}\left[\beta(-c_i^A)\sum\frac{\left(1-pre_{it}^A - rcv_{it}^A\right)}{(1+r)^t}\right]
\end{aligned} \tag{33}$$

Note that $1 - pre_{it}^A - rcv_{it}^A = ill_{it}^A + lyl_{it}^A$.

Figure 1: A nonfatal illness with recovery, but one that reduces life expectancy

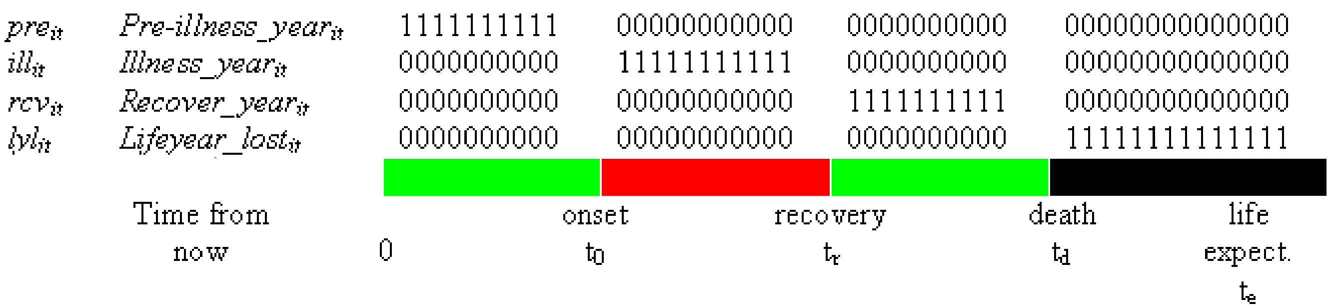

Figure 1:

Figure 2: A fatal illness

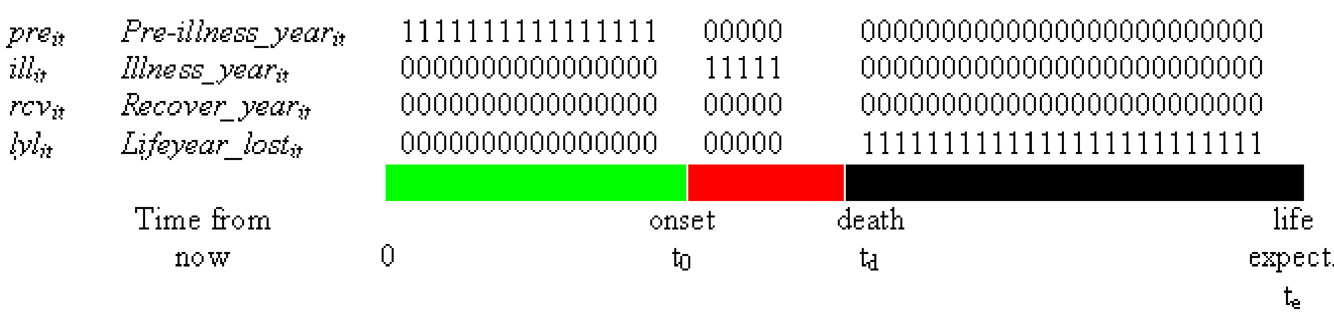

Figure 2:

Figure 3: A nonfatal illness with recovery, but one that reduces life expectancy, death event explicit

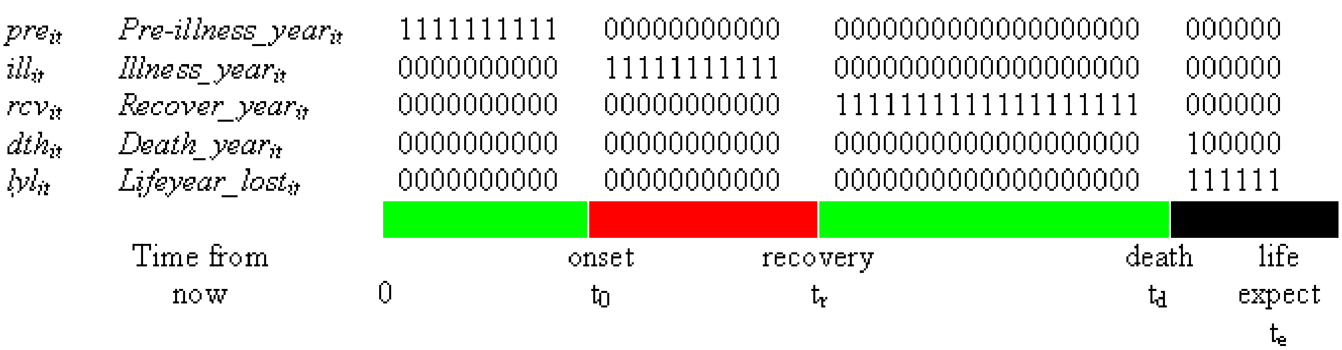

Figure 3: