Willingness to Pay for Health Risk Reductions: Differences by Type of Illness

Working Paper: Preliminary and Incomplete

Trudy Ann Cameron
Department of Economics
University of Oregon

J.R. DeShazo
UCLA Lewis Center for Regional Policy Studies
University of California, Los Angeles

Erica H. Johnson*
School of Business Administration
Gonzaga University
502 E. Boone Ave., AD Box 9
Spokane, WA 99258
johnsone@jepson.gonzaga.edu

June 2009

Keywords: value of a statistical life (VSL), value of a statistical illness profile, stated preference, willingness to pay, health risk reduction, health threat

JEL Classifications: Q51, I18

Acknowledgements: This research has been supported by a grant from the National Science Foundation (SES-0551009) to the University of Oregon (PI: Trudy Ann Cameron). It employs original survey data from an earlier project supported by the US Environmental Protection Agency (R829485) and Health Canada (Contract H5431-010041/001/SS). Additional support has been provided by the Raymond F. Mikesell Foundation at the University of Oregon. Office of Human Subjects Compliance approval filed as Protocol #C4-380-07F at the University of Oregon. This work has not been formally reviewed by any of the sponsoring agencies. Any remaining errors are our own.

**Corresponding Author*

Willingness to Pay for Health Risk Reductions: Differences by Type of Illness

ABSTRACT

In this paper, we examine how individual willingness to pay (*WTP*) for health risk reductions varies with the type of health threat in question. Our research focuses on systematic differences in *WTP* for health risk reductions across different types of major health threats, including five types of cancers (breast cancer, prostate cancer, colon cancer, lung cancer, skin cancer), chronic heart disease (as well as sudden heart attacks), respiratory disease, strokes, diabetes, Alzheimer's disease and traffic accidents. Our empirical results suggest that the marginal disutility from each type of health state differs across categories of illness (or injury). This suggests that models which constrain the estimated marginal utility parameters for different health states to be the same across all illnesses may be too restrictive and may cause us to lose information that may be very valuable from a policy perspective. The main contribution of this paper is to reinforce the case for why it may not be reasonable to assume one common value for *WTP* for risk reductions across all types of health risks.

1. Introduction

When it is necessary to conduct a formal benefit-cost analysis for a policy which protects human health or human lives, government agencies typically monetize the benefits from health risk reductions using a measure known as the Value of a Statistical Life (*VSL*). The *VSL* is an ex ante measure of individuals' willingness to pay (*WTP*) to achieve a change in the risk of premature death. The same *VSL* estimate is typically used whether the program in question leads to a reduction in the risk of lung cancer, heart disease, or any other illness. The U.S. Environmental Protection Agency (EPA) uses a *VSL* on the order of $6-7 million (2003 $) in its policy evaluations. In contrast, the Department of Transportation uses a *VSL* closer to $3-4 million, where the primary cause of death under consideration is often motor-vehicle accidents. The contribution of this paper is to demonstrate, within the context of a single study, that willingness to pay for health risk reductions varies systematically and significantly by disease type. *WTP* is simply a measure of inverse demand, and reductions in different types of health risks constitute differentiated products. It should not be surprising that similar risk reductions for different types of illnesses and injuries should be valued differently. Our results suggest that the EPA and other government agencies may wish to consider using different *VSL* values for different types of health threats to accurately represent more accurately the *WTP* of the population affected by a particular policy.

In order to average the empirical evidence of these tradeoffs across a variety of empirical studies which have considered different-sized risk reductions, it is conventional practice to scale each estimated *WTP* amount, proportionately, to a common (huge) risk change of 1.0. While it is tempting to interpret this scaled amount (the *VSL*) for a 1.0 risk change as "*WTP* to avoid certain death," this is inappropriate. The evidence concerning *WTP* measures is all based on very tiny

risk differences. The *VSL* value that reflects the average relationship between WTP and risk changes is never used to place a dollar value on prevention of one particular person's death with certainty. Instead, the *VSL* is scaled back down to its implied *WTP* for some typically tiny risk change that would be produced by a given policy, and the corresponding small *WTP* amount is summed across the affected population.

There are two typical ways of estimating a *VSL*: revealed preference methods, most commonly wage-risk studies, and stated preference methods, mainly contingent valuation and conjoint choice experiments. Typical *VSL* estimates from all of these methods range between about $1 million and $10 million. Wage-risk studies are the most common way of estimating a *VSL*. The methodology uses a hedonic regression that uses employee-specific characteristics, job characteristics and the actuarial job fatality risk as the explanatory variables. The method is described in an overview of revealed preference *VSL* estimates in Aldy and Viscusi (2007). This estimation strategy allows an estimate of the trade-off between risks and wages made in the labor market and provides a *WTP* to reduce mortality risk. The average participant in wage-risk studies is 36-37 years old and the risk involved in this decision is typically a sudden, accidental death that occurs at the workplace (Robinson (2007)).

Unfortunately for environmental economists, this measure of a *VSL* may not be a good fit when it is "transferred" to the population affected by many environmental policies. Children, the elderly and the infirm, who are not generally represented by a sample of individuals from the active labor force, are likely to enjoy a large share of the benefits from improved environmental quality. The trade-offs between risk and income evidenced within the working-age population may not match the trade-offs willingly made by members of the groups most affected by impaired environmental quality. In addition to the differences between the relevant populations,

the types of health risks from environmental exposures are different from risks faced in the labor market. Often the risks from poor environmental quality do not cause sudden death, but instead may have a long latency period before an individual develops symptoms.

Other examples where a *VSL* has been estimated via revealed preference data include analysis of individuals' willingness to exceed mandated speed limits, (Ashenfelter and Greenstone (2004)), and through purchases of safety equipment (Hakes and Viscusi (2007) and Atkinson and Halvorsen (1990)).

Stated preference surveys are another tool for measuring *VSL*. Surveys can be targeted to represent the general population or a sub-population may be identified for which individuals are particularly affected by a specific health threat. There are different types of stated preference survey formats. In contingent valuation methods, preferences tend to be elicited for a particular program consisting of a fixed bundle of attributes at varying prices. In conjoint choice experiments, choice sets are often designed to ensure that all specified attributes of the alternatives, including price, are varied independently. The main characteristics of a good stated preference study, according to Krupnick (2007), includes having a large sample size, passing scope and construct validity tests, and asking debriefing questions in the survey to identify potential problems such as scenario rejection.

Although most researchers would prefer to use revealed preference data in assessments of willingness to pay, such preference data are not always available. In the health risk reduction context, it will typically be necessary to control for differences in disease latency (i.e. time-to-onset), the duration of the illness or injury, whether or not the affliction is fatal, and the number of life-years that may be lost as a result of having suffered from this illness or injury. Hedonic wage studies may be good at capturing current period actuarial risks of sudden death, but they

are ill-suited to modeling illnesses which may have long latency periods, and those which may be chronic—namely, with long periods of pre-mortality morbidity.

The standard one-size-fits-all *VSL* is imperfect for various reasons. It is convenient to have a standardized measure of an inverse demand (a *WTP*) for health risk reductions. However, a constant *VSL* implies that the demand does not vary with individual characteristics like the demand for most commodities would. It is simply a fixed value and does not vary with income, gender, or household characteristics such as the presence of young children. Neither does it vary with the type of illness or accident which is addressed by the risk reduction policy in question.

Critics of the constant *VSL* also point out that it fails to acknowledge any difference according to age—it attaches the same value to premature death by one year as it does to premature death by eighty years. Consequently, a construct called the Value of a Statistical Life-Year (*VSLY*) is sometimes used to quantify *WTP* in addition to or in place of the *VSL*. The *VSLY* is calculated by dividing the *VSL* by average remaining life expectancy. A typical *VSLY* thus assumes that each year is equally valuable. Much empirical work has been done recently to show that the *VSL* does vary by individual characteristics, such as age, (e.g. Aldy and Viscusi (2007), Krupnick (2007) and Robinson (2007)). There is still much research to be done to explain how *VSL* varies with income, gender, life expectancy, and current health status as described in Hammitt (2007). This present paper attempts to tackle one more aspect of how *WTP* varies across health risks by looking at *WTP* to reduce the risks of suffering specific diseases.

In this paper—based upon a large general-population stated preference survey—we examine how individual willingness to pay (*WTP*) for health risk reductions varies systematically across different types of major health problems, including five types of cancers (breast cancer,

prostate cancer, colon cancer, lung cancer and skin cancer), chronic heart disease (as well as sudden heart attacks), respiratory disease, strokes, diabetes, Alzheimer's disease and traffic accidents. There are eleven different health threats total, including one gender-specific illness that is designated as breast cancer if the respondent is female, and prostate cancer if the respondent is male. In this conjoint choice study by Cameron and DeShazo (2008), respondents are shown five (independent) choice scenarios each involving three alternatives: two different risk-reduction programs and the status quo. For each respondent, ten of the eleven possible major health threats were used (two in each of the five choice sets).

We then estimate a choice model that allows us to simulate estimates of the *WTP* for microrisk reductions for an entire illness profile. An illness profile is a sequence of future health states including pre-illness years, sick years, potential recovered years, and lost-life years. This *WTP* measure is more general than the conventional *VSL*.

There are many reasons why people may be willing to pay different amounts to avoid similar illness profiles attributed to different types of illness or injury. Health risks may differ in terms of the degree of dread associated with them or with the perceived controllability of the risk (Slovic (1987)). There may be a "cancer premium" as noted in Savage (1993). Van Houtven et al. (2008) finds strong evidence for a "cancer premium" and finds that preferences differ with the length of the latency period. *WTP* by disease may also vary with the respondent's personal beliefs concerning their individual subjective risk of contracting a disease, their current health-related behaviors, and even with the extent to which they may feel that they have room to improve their health-related behaviors in order to reduce the risks of different diseases.

Savage (1993) finds that *WTP* differs across four specific causes of death: stomach cancer, plane crash, automobile crash and a home fire. *WTP* is positively related to the perceived

risk of exposure to that particular cause of death and the extent to which they dread the cause, and negatively related to the amount respondents know about that particular cause of death. He also finds that the *WTP* for stomach cancer is more than double the *WTP* for the other causes, which he attributes to the lack of familiarity with stomach cancer and the dread associated with it. Savage (1991) finds similar responses to dread in a previous study about nuclear power plant accidents.[1]

Sunstein (1997) investigates people's assessments of *WTP* for "bad deaths"—deaths that seem unusually horrible for some reason. There are many reasons that the cause of death may be perceived as "bad." The cause of death may seem less controllable, may affect certain demographic groups more than others, and may cause unusually long and severe bouts of suffering (i.e. pre-mortality morbidity). In a study of 116 University of Chicago law students, Sunstein found that about 40% of respondents felt that an avoided death from cancer was worth more than one avoided death from a heart attack. Although respondents may harbor more dread concerning some diseases than others, Sunstein concludes that the major policy considerations need to be the number of lives saved, the numbers of life-years saved, the quality of life during those saved years, and the cost-effectiveness of the programs (i.e. he does not advocate making policy simply based on the amount of dread associated with a disease).

Hammitt and Liu (2004) study systematic differences in *WTP* based on a latency period and the amount of dread associated with the disease. They find *WTP* estimates are higher to reduce the risk of cancer than to reduce the risk for other similar chronic diseases. They also find that *WTP* estimates may vary greatly between *WTP* to avoid a fatal accident and to avoid a disease that comes after a latency period or one that involves a large amount of dread. This

[1] For more articles on cancer, risks, and dread, see Trumbo et al. (2007), Chilton et al. (2002) and Chilton et al. (2006).

suggests that typical *VSL* estimates from wage-risk studies, for example, may not be ideal for valuing the benefits due to environmental policies.

Other researchers have found various results relating to specific health threats. Subramanian and Cropper (2000) compare environmental and public health programs and find that the seriousness of the risk and the number of lives saved matters for the rate of substitution between environmental and public health programs, but that the lack of controllability and the involuntary nature of environmental risks does not matter. Alberini et al. (2004) find that individuals who already have a disease appear not to have a lower overall *WTP* than those who do not. Vassanadumrongdee and Matsuoka (2005) find that perceived differences in dread, severity, controllability, and personal exposure between air pollution and traffic accidents have little effect on the *VSL*.

In this paper, we contribute the literature by looking at systematic differences in *WTP* for measures to reduce an individual's risk of suffering from one of eleven major health threats. The paper is structured as follows. Section 2 describes the available survey data. Section 3 describes our random utility choice model based on discounted expected utility and Section 4 describes its empirical implementation. Section 5 reviews our estimation results and Section 6 covers the implications of our estimated models for estimates of *WTP* which generalize conventional *VSL* estimates. Section 7 concludes, with a number of relevant caveats.

2. Available Data

The stated preference dataset from Cameron and DeShazo (2006) provides sufficient information to permit an analysis of differences in *WTP* by type of disease. The survey was administered by Knowledge Networks, Inc. to a random sample of respondents in the United States. Respondents are members of randomly selected households in the United States who are

offered free internet access in return for completing a few surveys every month. Since these respondents are part of a standing consumer panel, a large quantity of demographic and background information, such as health status and health history, is available for every member of the panel.

The survey has five parts.[2] The first part asks respondents about their personal health profile and their subjective risks of getting these diseases. The second part is a risk tutorial where risks are displayed in three different ways and respondents are required to answer risk comprehension questions. After thorough preparation, the third part of the survey asks the respondent to consider five different three-alternative conjoint choice sets. In each choice scenario, respondents choose between Program A, Program B, and the status quo (neither program) as seen in the sample choice matrix in Figure 1. Each program reduces the risk that the individual will suffer a particular illness profile. The health risk reduction programs, as described to respondents, consist of a diagnostic pin-prick blood test given by the individual's doctor once per year that indicates whether the individual is at risk for the illness. If the blood test indicates the individual is at risk, then the doctor would prescribe medication and life-style changes (such as diet and exercise) and continue to monitor the individual.[3] Each illness profile consists of a brief description which includes the age of the individual when the illness starts, the duration of the illness, the symptoms and treatments, and the expected effects on life expectancy. The risk reduction programs are characterized in terms of the expected risk reduction achieved by the

[2] For more detail, see an annotated survey at:
http://www.uoregon.edu/~cameron/vsl/Annotated_survey_DeShazo_Cameron.pdf

[3] For traffic accidents, the program is described as "new airbag, braking, and impact reduction technologies that are becoming available. These will reduce your chance of injury or death due to auto accidents. These technologies can be built into new vehicles, or added to existing vehicles. You will probably pay the cost of these technologies all at once when you buy a new car or have the equipment installed in an older one. When we describe costs, we will convert them to monthly costs and also annual costs to make them easier to compare across programs."

program, and the cost of the program (expressed in both monthly and annual terms). All of the attributes are randomized, subject to basic plausibility constraints. The fourth part of the survey consists of debriefing questions which follow up on each conjoint choice task. The fifth part of the survey was taken separately by all panelists and gathers socio-demographic information that can be readily merged with the data collected expressly for this survey.

The survey was administered to 2,439 respondents with a 79% response rate among invited panelists. In certain cases, a respondent or a specific choice set was dropped from the estimating sample. Respondents were excluded from the estimating sample if they failed to pass skill-testing questions about risk comprehension, if they rejected outright the types of choice scenarios the survey posed, or because of a small error in the randomized design of the survey.[4] After these data exclusion criteria are applied, 1,619 respondents remain. These respondents considered 7,520 choice sets involving 13,696 risk reduction programs (and a total of 20,544 alternatives when the status quo is included). Descriptive statistics are shown in Table 1.

3. Utility-Theoretic Choice Model

Survey respondents choose from three alternatives in each choice set. There are two risk reduction programs, Program A and Program B, and the status quo (Neither Program), which are denoted A, B and N. Each program reduces the risk of facing an illness profile attributed to one of eleven different "illness labels" but each program has an associated monetary cost. The program cost is assumed to apply only during pre-illness years and recovered years, so the individual would not pay for the program if he or she were to fall ill as described in the illness profile. An illness profile is a sequence of future health states that includes a specified

[4] Due to a lack of risk comprehension 1,236 choices (4,887 alternatives) were dropped, due to scenario rejection (where the respondent only chose scenario rejection as the reason for choosing the "neither program" alternative) 2,236 choices (6,708 alternatives) were dropped, and due to an error in the randomization of the survey 332 choices (996 alternatives) were dropped.

combination of pre-illness years, sick years, post-illness (recovered) years and lost-life years. Respondents are assumed to choose the alternative that gives them the highest level of utility.

This utility-theoretic choice model is described in detail in Cameron and DeShazo (2006), but we offer a brief explanation of the model in this paper. For simplicity, consider just the pair-wise choice between Program A and Neither Program.[5] We assume that the utility of an individual, i, at time, t, depends upon net income in that period, Y_{it} minus the cost of any program, and the health state they experience in that period. The survey considers only single spells of any given illness. In any given period, the individual will be in one of four possible health states. These are recorded via four indicator variables: $1(pre_{it})$ for pre-illness years, $1(ill_{it})$ for illness-years, $1(rcv_{it})$ for recovered or post-illness years, and $1(lyl_{it})$ for lost-life years. We can write the individual's indirect utility function in each time period, t, as:

$$V_{it} = \beta_0 Y_{it} + \beta_1 Y_{it}^2 + \alpha_0 1(pre_{it}) + \alpha_1 1(ill_{it}) + \alpha_2 1(rcv_{it}) + \alpha_3 1(lyl_{it}) + \eta_{it} \qquad (1)$$

The health states are mutually exclusive and they are also exhaustive, meaning that the individual experiences one, and only one, of these four health states at a time.

There is uncertainty about whether the individual will actually fall sick from the disease, so we model each choice as depending upon expected indirect utility, with the expectation taken across the sick (S) and healthy (H) outcomes. Participation in Program A instead of the status quo is described as altering the probability of getting sick from Π_i^{NS} to Π_i^{AS}. Furthermore, each illness profile extends through the remainder of the individual's life expectancy, so we discount future time periods using discount rate r and discount factor $\delta^t = (1+r)^{-t}$ to get the indirect utility in terms of present value, which we denote as PDV. The individual is assumed to choose

[5] The three-way choice between two programs and neither program is analogous.

Program A over the status quo alternative (Program N) if his or her discounted expected utility is greater under Program A:

$$PDV\left(\Pi_i^{AS} V_i^{AS} + \left(1-\Pi_i^{AS}\right) V_i^{AH}\right) - PDV\left(\Pi_i^{NS} V_i^{NS} + \left(1-\Pi_i^{NS}\right) V_i^{NH}\right) > 0 \tag{2}$$

The present discounted number of years making up the remainder of the individual's nominal life expectancy, T_i, is given by $pdvc_i^A = \sum_{t=1}^{T_i} \delta^t$. Other relevant discounted spells, also summed from $t=1$ to $t=T_i$ include $pdve_i^A = \sum \delta^t 1\left(pre_{it}^A\right)$, $pdvi_i^A = \sum \delta^t 1\left(ill_{it}^A\right)$, $pdvr_i^A = \sum \delta^t 1\left(rcv_{it}^A\right)$, and $pdvl_i^A = \sum \delta^t 1\left(lyl_{it}^A\right)$. Since the different health states exhaust the individual's nominal life expectancy, $pdve_i^A + pdvi_i^A + pdvr_i^A + pdvl_i^A = pdvc_i^A$. Finally, to accommodate the assumption that each individual expects to pay program costs only during the pre-illness or recovered post-illness periods, $pdvp_i^A = pdve_i^A + pdvr_i^A$, is defined as the present discounted time over which payments must be made.

To further simplify notation, let $cterm_i^A = \left[\left(1-\Pi_i^{AS}\right)\right] pdvc_i^A + \Pi_i^{AS} pdvp_i^A$. Let $yterm_i^A = \left[-pdvc_i^A + \Pi_i^{AS} pdvi_i^A + \Pi_i^{NS} pdvl_i^A\right]$ and $pterm_i^A = \Pi_i^{AS}\left[\alpha_1 pdvi_i^A + \alpha_2 pdvr_i^A + \alpha_3 pdvl_i^A\right]$. Then the expected utility-difference that drives the individual's choice between Program A and the status quo can then be defined as follows (where there will be an analogous term for the utility difference between Program B and the status quo in our three-alternative model):

$$\begin{aligned}\Delta PDV\left(E_{S,H}\left[V_i\right]\right) = {} & \beta_0\left\{\left(Y_i - c_i^A\right) cterm_i^A + Y_i\, yterm_i^A\right\} \\ & + \beta_1\left\{\left(Y_i - c_i^A\right)^2 cterm_i^A + Y_i^2\, yterm_i^A\right\} \\ & + \alpha_1\left\{\Delta\Pi_i^{AS}\, pdvi_i^A\right\} + \alpha_2\left\{\Delta\Pi_i^{AS}\, pdvr_i^A\right\} + \alpha_3\left\{\Delta\Pi_i^{AS}\, pdvl_i^A\right\} + \varepsilon_i^A\end{aligned} \tag{3}$$

The option price, in the sense of Graham (1981), is the common certain maximum payment that makes an individual indifferent between paying for the program and having the risk reduction, or not paying for the program and not having the risk reduction. In the context of the detailed model from Cameron and DeShazo (2008), we can solve the discounted expected indirect utility-difference for the value of this common certain payment:

$$\hat{c}_i^A = Y_i - f^{-1}\left(\frac{\left(\beta_0 + \beta_1 Y_i\right) yterm_i^A + pterm_i^A + \varepsilon_i^A}{-\left(\beta_0 + \beta_1 Y_i\right) cterm_i^A}\right) \tag{4}$$

where $f(Y) = \left(\beta_0 + \beta_1 Y_i\right) Y_i = \beta_0 Y_i + \beta_1 Y_i^2$ and $f^{-1}(\cdot)$ is the solution to a quadratic form.

The expected present value of the certain payment can then be calculated for the individual's remaining lifetime and can be written as:

$$E_{S,H}\left[PV\left(\hat{c}_i^A\right)\right] = cterm_i^A\left[\hat{c}_i^A\right] \tag{5}$$

Next, we normalize this expected present value of the certain payment by a risk change of one-in-one-million in order to allow comparison between different risk changes. We divide $E_{S,H}\left[PV\left(\hat{c}_i^A\right)\right]$ by the absolute size of the risk reduction and multiply by .000001 to get:

$$WTP = \frac{E_{S,H}\left[PV\left(\hat{c}_i^A\right)\right]}{\Delta\Pi_i^A} \times .000001 \tag{6}$$

The *WTP* is a marginal rate of substitution (with the marginal utility of the sequence of health states in the numerator and the marginal utility of income in the denominator) for a microrisk reduction. Since the marginal utility of an adverse illness profile is in the numerator of the *WTP*, an increase in the marginal (dis)utility of any component of an illness/injury profile of health states (illness years, recovered years, and lost life-years) will increase the *WTP*. Since the marginal utility of income is in the denominator, an increase in the marginal utility of income will decrease the *WTP*.

In order to get a simulated *WTP* using our choice data, we need an approximate joint distribution for the illness profile of a particular illness, which comes from epidemiological studies. We also need a joint distribution of age, gender, and income level. Then, we make a large number of draws from these two joint distributions and simulate the *WTP* values. The mean of the distribution of *WTP* estimates can be interpreted as the model's prediction of the average *WTP* of this particular illness and for this particular population.

As described in Cameron and DeShazo (2008), the data suggest that the basic five-parameter, homogeneous-preferences model given in equation (3) is dominated by a specification that is not merely linear in the terms involving present discounted health-state years. Rewriting the final term in equation (3) gives:

$$\begin{aligned} &\alpha_1\left\{\Delta\Pi_i^{jS}\, pdvi_i^j\right\}+\alpha_2\left\{\Delta\Pi_i^{jS}\, pdvr_i^j\right\}+\alpha_3\left\{\Delta\Pi_i^{jS}\, pdvl_i^j\right\} \\ &\quad=\Delta\Pi_i^{jS}\left[\alpha_1 pdvi_i^j+\alpha_2 pdvr_i^j+\alpha_3 pdvl_i^j\right] \end{aligned} \tag{7}$$

Where $j = A, B, N$, and $pdvX_i^N = 0$ for $X = i, r, l$. This simple linear specification fails to explain respondents' observed choices as well as a model that employs shifted *logarithms* of the $pdvX_i^j$ terms. Starting from a form that is fully translog (including all squares and pairwise interaction terms for the three log terms), and retaining only those terms where the α coefficients are statistically significantly different from zero, this final term becomes:

$$\Delta\Pi_i^{AS}\left[\begin{array}{l}\alpha_1\log\left(pdvi_i^A+1\right)+\alpha_2\log\left(pdvr_i^A+1\right)+\alpha_3\log\left(pdvl_i^A+1\right) \\ +\alpha_4\left\{\log\left(pdvl_i^A+1\right)\right\}^2+\alpha_5\left\{\log\left(pdvi_i^A+1\right)\log\left(pdvl_i^A+1\right)\right\}\end{array}\right] \tag{8}$$

Finally, because the opportunity for longer durations in each health state is correlated with the youth of the respondent, the α coefficients must be allowed to differ systematically with the respondent's current age wherever this generalization is warranted by the data. This

leads to a model where $\alpha_3 = \alpha_{30} + \alpha_{31} age_i + \alpha_{31} age_i^2$, and analogously for α_4 and α_5. This quadratic-in-age systematic variation in parameters permits non-constant age profiles for the model's *WTP* estimates, and this sample tends to produce the usual higher values during middle age and lower values for younger and older respondents. One final parameter stems from a correction for differing sampling propensities, which appears to be relevant only for the $\log\left(pdvi_i^A + 1 \right)$ term.

4. Empirical Specification

In this paper, we build on this basic thirteen-parameter specification described in Cameron and DeShazo (2008) by generalizing the marginal utility parameters so that they vary systematically by disease type. There is no theory to recommend specific functional forms whereby these individual parameters should be expected to vary, so we introduce disease types and their interactions with other relevant individual attitudes and characteristics in an effort to build an understanding of how the implicit information contained in the disease labels can influence respondents' stated preferences over the alternative health-risk reduction programs proposed in our survey.

All of the specifications we consider retain the two basic "income" terms in the original thirteen-parameter model. The two parameters to be estimated correspond to the linear- and quadratic-in-net income terms in the underlying assumed indirect utility function. However, because the time profile of income and program costs will depend on the sequence of health states if the individual suffers from the disease in question, it is necessary to construct the present discounted expected net income terms under each risk-reduction program and under the status quo. Fortunately, the underlying linear- and quadratic-term coefficients persist as modifiers in the constructed discounted expected net income variables, which we identify in the tables as

"linear net income term" and "quadratic net income term." The estimated coefficient on the linear term has the expected positive sign and the sign for the coefficient on the quadratic term (which is present to allow for diminishing marginal utility of net income) has the expected negative sign in all models. The income terms in the model enter into the denominator of the *WTP* formula for health risk reductions (which is a nonlinear function of the estimated parameters).

In the analyses described in this paper, we maintain the assumption that the individual's marginal utility of net income is unaffected by any implicit or imputed characteristics conveyed by illness label used for each program. Differentiating each parameter by type of illness expands the number of parameters by a factor of twelve. There is no evidence of any heterogeneity by disease type in the coefficient on the quadratic-in-net-income term, although in simple models there is some suggestion that the linear coefficient in the marginal utility of income may be lower when the illness in the program is described as being respiratory disease. In richer models, however, that effect disappears and a different one, for diabetes, materializes. As usual, it is difficult to allow every parameter in a model to vary systematically with the same long list of shifters. The question of where to introduce the heterogeneity appears to hinge on which placement seems to produce the most robust results.

Our focus in this paper is the relevance of any other unspecified implicit attributes of each illness—conveyed by the arbitrarily assigned illness labels, yet not captured merely by the time periods in different health states (latency, sick-years, and lost life-years). It seems plausible that the effects of illness labels could act by shifting the marginal ex ante (dis)utility of sick-years from that affliction, or even the marginal ex ante (dis)utility of lost-life years. In the latter case, however, there is the possibility that "dead is dead," and that the affliction from which you

die has much more to do with the marginal (dis)utility of the sick-years leading up to death than it will with the disutility of being (prospectively) prematurely dead for some number of years. Alternatively, illness label effects could enter simply as a "lump" of additional indirect utility affecting preferences for each alternative in the choice set. The results from specifications that analyze each of these affects individually and the full model with all shifters are described in Appendix II, but we focus here only on the preferred parsimonious model.

In our specifications, we use heart disease as the base case and all other marginal utilities are differentials relative to the (dis)utility from a discounted sick-year with heart disease. It is worth emphasizing that the illness labels used in our survey were assigned randomly to different illness profiles. Some combinations were implausible (such as sudden death from diabetes or Alzheimer's disease), so these combinations were removed. However, it would be possible for two identical illness profiles to appear in the study, but with different illness labels. Thus we can be confident that the effects we find for the impacts of labels are not merely picking up attributes of the illness profiles.

In all models the illness effects on the marginal (dis)utility of sick-years are permitted to vary with five additional variables (as well as an interaction between them): confidence, vulnerability, controllability, subjective risk, and smoker.

4.1 Confidence

"Confidence" is the respondent's answer to the general question "Imagine you experience one of the major illnesses described in this survey. How confident are you that your diagnosis and treatment by your current health care provider would be both timely and of high quality?" Possible responses include -1= "not at all confident," 0= "somewhat confident," and "+1=highly confident." Note that we normalize each attitudinal variable on the median value in the sample

to facilitate interpretation of the base effect, which will apply for the individual with median values of these shifters. While we could allow for greater generality by capturing this factor with a pair of dummy variables, we treat it as an approximately cardinal variable to conserve on parameters. If a respondent does not have a high level of confidence in their access to, or quality of, health care, they may be more willing to pay for a preventative program.

4.2 Vulnerability

We also allow health state terms to vary by perceived health "vulnerability," which is the respondent's answer to the question "What is the chance that you will experience, either for the first time or as a recurrence, one of the major illnesses we discussed within the next 20 years?" The response options are coded as -2="very unlikely," -1="somewhat unlikely," 0="somewhat likely," and 1= "very likely." If some respondents feel that they have a higher chance of suffering from a major illness over this time horizon, we expect that their prospective (dis)utility will be greater and they will have a higher *WTP* for health risk reduction programs. Since the perception of health vulnerability is likely to be correlated with age, however, we allow the effect of perceived vulnerability to vary with the respondent's current age (and we also allow for age to directly shift the influence of each illness label on the marginal (dis)utility of a sick-year), Older respondents are more likely to face at least one of these named diseases in the next twenty years, so if "vulnerability" were alone in the model, it might merely be picking up this age effect. The relationship between age and subjective future health vulnerability is shown in Figure 2. Indeed, many of our respondents' experiences and perceptions of health problems vary markedly with the individual's current age. Age is likely to proxy for a number of factors which will affect the salience of specific illnesses.

4.3 Controllability

We also allow for heterogeneity in the marginal (dis)utility of sick-year terms with respect to the extent to which the individual feels the disease is controllable, and how much the individual feels he or she is at risk for the disease. "Controllability" is the respondent's answer to an auxiliary survey question worded as follows: "How much do you think that improving your lifestyle or habits would reduce your risk of [each class of health risk]." Response options ranged from -2= "very little," to +2= "a lot." Again, the variable is treated as approximately continuous and we allow it to enter linearly. (This variable has been normed on the "neutral" category.) The anticipated effect of this attitude on demand for health-risk reductions programs cannot be signed in advance. If respondents feel that a disease is more controllable, they might express a greater demand for a program to prevent it, since the program would be more likely to work.[6] On the other hand, if they feel the disease is controllable, perhaps they do not need a special prevention program since they feel they will be able to control the risk of the disease on their own.[7]

4.4 Subjective Risk

Finally, we expect that individuals who feel more at risk for getting a *particular* disease would have a higher *WTP* for a risk reduction program for that disease. Our survey asked respondents to "Think about your health, your family history, and hazards to which you are exposed. Which illnesses or injuries do you feel most at risk of experiencing over your lifetime?" Response options ranged from -2= "low risk," to +2= "high risk." We expect that for

[6] The health-risk reduction programs described in the survey involve a simple diagnostic test. Respondents are told: "If a test says that you have a problem, your doctor could prescribe medication and life-style changes that reduce your risk of getting the illness. You would continue to be monitored."

[7] It is not possible to interact the "controllability" and "subjective risk" variables with the disease indicators alone because they are specific to each disease.

illness labels corresponding to specific health threats for which the respondent feels particularly at risk should have a greater effect on the magnitudes of the coefficients.

4.5 Smoker

We allow interactions between the disease labels and confidence, vulnerability, age and age and vulnerability. We also control for whether individuals reveal that there is room for them to reduce their health risks by improving their lifestyle or habits if they quit smoking. We use this acknowledgement to identify each individual as a current smoker or non-smoker. We then allow for an interaction between current smoker and two of the health threats, respiratory disease and lung cancer. If the respondent could improve their health if they smoked less, which means they smoke, they may feel more vulnerable to experiencing respiratory disease and lung cancer.

We considered the possibilities that heterogeneity in demand by illness label may affect the marginal (dis)utility of discounted sick-years, the basic indirect utility associated with each program *regardless* of the time profile of the illness, and the marginal (dis)utility of discounted lost life-years. Results are shown in Appendix II, but we should note here that we find numerous differences by illness label in all three types of generalizations.

The overwhelming number of coefficients in the most extensive model needs to be reduced, so we focus on the results from preferred parsimonious model in the next section. Given the randomized design of the program attributes, there is much less risk that multicollinearity will complicate the process of identifying stable and statistically significant illness label effects by pruning away persistently insignificant variables. Of course, the variables which we use to control further, for respondent attitudes, are correlated to a certain extent, but we will be zeroing out the influence of these variables in our simulations. These are merely control variables, included where necessary to help us better discern the tendencies in the data

which can be identified for a respondent with "neutral" attitudes on all dimensions, for selected specific age levels.

5. Estimation Results

Table 2 shows selected coefficients on the basic terms, age, and smoker from the parsimonious version of the extensive model that retains the most robust and persistently statistically significant coefficients. The full parsimonious model, showing the full set of the included controls, is in Appendix I.

Our empirical results suggest that the disutility associated with the disease labels of heart attacks, breast cancer (for female respondents), prostate cancer (for male respondents), colon cancer, strokes, and diabetes is larger than the disutility associated with one year of heart disease. The disutility for the disease labels of lung cancer, skin cancer and respiratory disease appears to be less than the disutility associated with heart disease. For traffic accidents, diabetes and Alzheimer's disease, the disutility for these labels appears to decrease with the age of the respondents. Smokers appear to have a greater disutility from lung cancer and respiratory disease than non-smokers.

The estimated coefficients suggest suggest that the marginal (dis)utility of a sick-year for many diseases is statistically indistinguishable from that associated with heart disease. These results suggest that respondents view a sick-year due to heart attacks, prostate cancer (for male respondents), and strokes as statistically significantly different from heart disease.

Lost-life years due to Alzheimer's disease are statistically significantly different from lost-life years due to heart disease, but there appears to be more significance related to whether respondents think they will receive good medical treatment for that disease and how controllable it is.

The more interesting and intuitive results, that use these marginal utility parameters, are the *WTP* estimates that are discussed in the next section.

6. Implications for *WTP* difference by disease

We use this parsimonious model when simulating *WTP* values. We will focus on variation by disease label, but also by age, smoking status, and two types of illness profiles. We will conduct these simulations of *WTP* for an individual who is assumed to have median levels (i.e. zero values) of the other control variables we employ—namely, the variables for confidence in the quality of future health care, subjective vulnerability to major health problems in the upcoming twenty years, the subjective controllability of each type of illness and the subjective risk of suffering each type of illness.

Table 3 shows examples of the predictions of the parsimonious model concerning the *WTP* for a microrisk reduction. The numbers in this table can be interpreted as the fitted willingness to pay for a 1 in 1,000,000 reduction in the risk of the specified illness profile (including latency, sick-years, and lost-life years from a particular named illness) for an individual with an income of $42,000. These values are generated by drawing 1000 values from the asymptotically normal joint distribution of the maximum likelihood parameter estimates and using each set of parameters to calculate the *WTP* from the basic formula in terms of a specific set of values for the variables and that replication's set of parameters. Across this simulated distribution, we acknowledge that respondents were given no opportunity to express negative willingness to pay. All they could do was to "not choose" a less desirable program. Thus we convert all negative calculated values to zero and report the mean of the resulting distribution. We also report the 5^{th} and 95^{th} percentiles of each distribution.

Table 3 displays results for a 30-year-old, a 45-year-old, and a 60-year-old. Two different illness profiles are considered. The first illness profile is sudden death now (due to each of the health threats—which may be unlikely in some cases). The "sudden death" *WTP* estimates are most similar to a conventional *VSL*. We also offer another example of an illness profile, involving ten years of latency, five years of illness, followed by death.

For a benchmark individual, our highest estimates of willingness to pay to reduce the risk of illness appear to be for smokers to reduce the risk of lung cancer. Across our three representative age groups, *WTP* estimates for these individuals are highest for the 45-year-old, at just over $11.00. Non-smokers, on the other hand, seem to have relatively low willingness to pay to reduce their risks of lung cancer. The *WTP* is less than $1.00 under the sudden death scenario. Under the more realistic latent illness scenario, the *WTP* is higher, at $2.36 for the 30-year-old, but it falls dramatically with age, amounting to only $0.40 for the 60-year-old.

For smokers, there is also a dramatic differential in willingness to pay for reductions in the risk of respiratory disease. The *WTP* for a microrisk reduction in sudden death from respiratory disease is on the order of $5.00 to $6.30 for smokers. For non-smokers, it ranges from only $0.38 down to $0.03. The latent illness scenario, which is probably more plausible, produces even higher *WTP* estimates for smokers, but only for the 30-year-old is the non-smoker *WTP* greater than $1.00. The *WTP* appears to be dramatically lower for the 45-year-old and essentially zero for a non-smoker who is already 60 years old.

Among other cancers, both breast cancer and prostate cancer are of considerable concern to all age groups, whether there is a latency period or not. These *WTP* values, on the order of $4.00 to about $8.70, when scaled up by one million to be compared with a *VSL*, are in the ballpark of numbers used currently by the U.S. EPA and the Department of Transportation.

Willingness to pay to reduce colon cancer risk is somewhat lower, slightly over half as great for the sudden death scenarios and about three-quarters as large for the scenario with ten years of latency and five years of illness followed by death.

In contrast, measures to reduce the risk of skin cancer attract very little interest. Across our six age/latency cases, only the 30-year-old, in the latent case, has a *WTP* greater than $1.00. Sixty-year-olds have negligible willingness to pay to reduce their risks of skin cancer.

Among the non-cancer illnesses, the largest *WTP* amounts are associated with heart disease and heart attacks. The numbers tend to be virtually identical in the sudden death case, which is reassuring, because there is no requirement that the data yield identical values for these two different disease labels, but logically we might expect them to be similar. For the case with latency, the *WTP* estimates are about $1.00 higher for heart disease than for heart attacks, although the difference is smaller for the 60-year-old.

Cerebrovascular illness (stroke) is only somewhat less of a concern, with *WTP* amounts on the order of $6.40 for the 30-year-old down to $5.40 for the 60-year-old. With ten years of latency and 5 years of sick-time, however, the *WTP* is somewhat lower, and it drops to only $2.70 for the 60-year-old.

Despite the incidence of Type II diabetes increasing with age, as shown in Figure 7, the *WTP* for diabetes (in the improbable sudden death case) drops from $5.3 for the 30-year-old to only $300,000 for the 60-year-old. For the scenario with latency and five years of illness, the *WTP* is higher for the 30-year-old (at $7.15), but it drops to only $0.07 for the 60-year-old.

Willingness to incur costs to lower the risk of traffic accidents is surprisingly low. However, we suspect that our choice scenarios may have left some to be desired. Respondents were told that they could buy equipment with their new car, or retrofit an older vehicle at a

specified annualized cost. For most of the illness profiles, however, the injury was described as occurring more than seven years into the future. It is possible that many people, assuming that they would not own their current car for more than seven years, might have been reluctant to pay for these measures. It was difficult to get this illness scenario to conform to the others. There is no plausible age at which one might become more likely to suffer an accident. In fact, most people perceive themselves to be at relatively little risk of an accident and feel that traffic accident risks are beyond their control.

Despite the potential for "scenario rejection" with the traffic accident risk reduction programs, the implied *WTP* for sudden death in the current year is still over $1.00 for the 30-year-old, although it drops to only $.07 for the 60-year-old. Figure 8 reveals that older respondents report lesser experience with traffic accidents, either for themselves or among their family and friends. For the case with latency, however, the 30-year-old has a *WTP* of $2.69. If scaled to be comparable to a *VSL*, the ninety percent interval contains the $3-4 million *VSL* amount used by the Department of Transportation in 2003. Our data certainly support the notion that willingness to pay to reduce highway risks is less than that for heart disease, stroke, diabetes (at least among the young) and several types of cancers.

The only disease for which there is evidence of dramatic increases with age is Alzheimer's disease. The thirty-year-old and the 45-year-old are not willing to pay much at all to reduce their risks of Alzheimer's, either in the current year or ten years down the road. This coincides with the very much lower incidence of Alzheimer's among younger people. For the 60-year-old, however, willingness to pay to reduce Alzheimer's now is about $2.20. In the latent case for the 60-year-old, where Alzheimer's will not begin until they are 70 and they will not die until they are 75, the *WTP* is lower, at just over $1.00.

7. Caveats and Conclusions

Using a stated preference survey concerning willingness to pay for health risk reductions, we look at systematic differences by disease. We use a random utility model framework for this analysis and then allow the parameters to shift with disease labels in our empirical analysis. We find systematic differences due to the disease labels by themselves, but also a number of differences, by disease, in the marginal (dis)utilities associated with prospective future adverse health states (sick-years and lost life-years).

Reductions in the risk of breast and prostate cancers, especially in the near term, seem to be valued even somewhat above the *VSL* currently employed by the U.S. EPA. Values for colon cancer reductions are somewhat lower, but the range of simulated values includes the roughly $6-7 million *VSL* used by the EPA. Reductions in lung cancer risks are of much lesser concern—except to smokers, where the *WTP* is on the order of $11.00 in the near term and even higher if some latency is involved. Non-smokers care relatively little about reducing lung cancer risks, and nobody seems to care very much about reducing skin cancer risks.

Willingness to pay to reduce risks from heart disease and heart attacks are very similar to each other (and to breast cancer and prostate cancer) , the *WTP* to reduce microrisks are on the order of $7.00 to $8.00, while strokes may be of somewhat lesser concern, perhaps similar to colon cancer (at least when some latency is involved) and in line with current EPA numbers.

Smokers appear to be as much concerned about reducing their risk of respiratory disease as men are about reducing prostate cancer risks if the risk involves some latency (but somewhat less so in the case of sudden death in the current period). The *WTP* values for a microrisk reduction are on the order of $4.00 - $8.00. Non-smokers, however, have very little interest in paying to reduce their risks of respiratory disease.

Diabetes is more of a concern among the young than among older people, whereas the reverse is true for Alzheimer's disease. Traffic accidents are of surprisingly little concern among older people, perhaps because they see themselves to be less at risk because they spend less time on the road, or because they believe themselves to be safer drivers. Reports of traffic accidents for respondents themselves, or among their family and friends, seem to decline with age as shown in Figure 8. Only for the youngest group does there appear to be a willingness to pay to reduce serious traffic accidents that, when scaled, approaches the *VSL* used by the Department of Transportation in 2003.

Concerning environmental threats to health, one might think first of respiratory disease and lung cancer from "criteria" pollutants and toxic air pollutants. Some portion of the population may also be aware of the role of air pollution in heart disease. Our results suggest that there may be a huge difference between the smoking and non-smoking populations in demands for health risk reductions via reductions in air pollution. We plan to explore further whether there is a difference according to the actual death rates from lung cancer and respiratory disease in the individual's county of residence (as a way to capture the influence of existing stressors in the form of air pollution). We may be able to recover from our database sufficient information to place respondents inside or outside an urban area, which may also help explain systematic differences in *WTP* for air pollution reductions.

One of the first questions in our survey elicited, from each respondent, his or her subjective opinion about the degree to which air quality and drinking water quality posed a threat to their health. We have yet to fully exploit this information to help explain variations in *WTP* by disease.

Overall, the substantial differences in willingness to pay for health risk reductions by type of illness or injury suggest that models which constrain the estimated marginal utility parameters for different health states to be the same across all illnesses may be too restrictive and may cause the loss of information that may be very valuable from a policy perspective. The differences in *WTP* estimates suggest that respondents have different *WTP* values for avoiding different diseases. This suggests that the types of health threats to be targeted by a specific policy (as well as the age groups the policy will affect) might be taken into consideration in benefit-cost analysis.

References

Alberini, A., Cropper, M., Krupnick, A., Simon, N.B. (2004). "Does the value of a statistical life vary with age and health status? Evidence from the US and Canada", *Journal of Environmental Economics and Management* 48, 769-792.

Aldy, J.E., Viscusi, W.K. (2007). "Age differences in the value of statistical life: Revealed preference evidence", *Review of Environmental Economics and Policy* 1, 241-260.

Ashenfelter, O., Greenstone, M. (2004). "Using mandated speed limits to measure the value of a statistical life", *Journal of Political Economy* 112, S226-S267.

Atkinson, S.E., Halvorsen, R. (1990). "The valuation of risks to life - evidence from the market for automobiles", *Review of Economics and Statistics* 72, 133-136.

Cameron, T.A., DeShazo, J.R. (2006). "A generalized model of demand for risk reductions: Estimating the value of a statistical illness profile", Department of Economics, University of Oregon Working Paper.

Cameron, T.A., DeShazo, J.R. (2008). "A generalized model of demand for risk reductions: Estimating the value of a statistical illness profile", Department of Economics, University of Oregon Working Paper.

Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., Spencer, A. (2002). "Public perceptions of risk and preference-based values of safety", *Journal of Risk and Uncertainty* 25, 211-232.

Chilton, S., Jones-Lee, M., Kiraly, F., Metcalf, H., Pang, W. (2006). "Dread risks", *Journal of Risk and Uncertainty* 33, 165-182.

Graham, D.A. (1981). "Cost-benefit-analysis under uncertainty", *American Economic Review* 71, 715-725.

Hakes, J.K., Viscusi, W.K. (2007). "Automobile seatbelt usage and the value of statistical life", *Southern Economic Journal* 73, 659-676.

Hammitt, J.K. (2007). "Valuing changes in mortality risk: Lives saved versus life years saved", *Review of Environmental Economics and Policy* 1, 228-240.

Hammitt, J.K., Liu, J.T. (2004). "Effects of disease type and latency on the value of mortality risk", *Journal of Risk and Uncertainty* 28, 73-95.

Krupnick, A. (2007). "Mortality-risk valuation and age: Stated preference evidence", *Review of Environmental Economics and Policy* 1, 261-282.

Robinson, L.A. (2007). "How US government agencies value mortality risk reductions", *Review of Environmental Economics and Policy* 1, 283-299.

Savage, I. (1991). "Psychological features affecting valuation of life", *Economics Letters* 35, 379-383.

Savage, I. (1993). "An empirical-investigation into the effect of psychological perceptions on the willingness-to-pay to reduce risk", *Journal of Risk and Uncertainty* 6, 75-90.

Slovic, P. (1987). "Perception of risk", *Science* 236, 280-285.

Subramanian, U., Cropper, M. (2000). "Public choices between life saving programs: The tradeoff between qualitative factors and lives saved", *Journal of Risk and Uncertainty* 21, 117-149.

Sunstein, C.R. (1997). "Bad deaths", *Journal of Risk and Uncertainty* 14, 259-282.

Trumbo, C.W., McComas, K.A., Kannaovakun, P. (2007). "Cancer anxiety and the perception of risk in alarmed communities", *Risk Analysis* 27, 337-350.

Van Houtven, G., Sullivan, M.B., Dockins, C. (2008). "Cancer premiums and latency effects: A risk tradeoff approach for valuing reductions in fatal cancer risks", *Journal of Risk and Uncertainty* 36, 179-199.

Vassanadumrongdee, S., Matsuoka, S. (2005). "Risk perceptions and value of a statistical life for air pollution and traffic accidents: Evidence from bangkok, thailand", *Journal of Risk and Uncertainty* 30, 261-287.

Figure 1: One example of a randomized choice scenario

Choose the program that reduces the illness that you most want to avoid. But think carefully about whether the costs are too high for you. If both programs are too expensive, then choose Neither Program.

If you choose "neither program", remember that you could die early from a number of causes, including the ones described below.

	Program A for Heart Disease	**Program B for Colon Cancer**
Symptoms/ Treatment	Get sick when 71 years old 2 weeks of hospitalization No surgery Moderate pain for remaining life	Get sick when 68 years old 1 month of hospitalization Major surgery Severe pain for 18 months Moderate Pain for 2 years
Recovery/ Life expectancy	Chronic heart condition Die at 79	Recover at 71 Die of something else at 73
Risk Reduction	5% From 40 in 1,000 to 38 in 1,000	50% From 4 in 1,000 to 2 in 1,000
Costs to you	$15 per month [= $180 per year]	$4 per month [= $48 per year]
Your choice	☐ **Reduce my chance of heart disease**	☐ **Reduce my chance of colon cancer**
	☐ **Neither Program**	

Figure 2: Confidence in quality of medical care

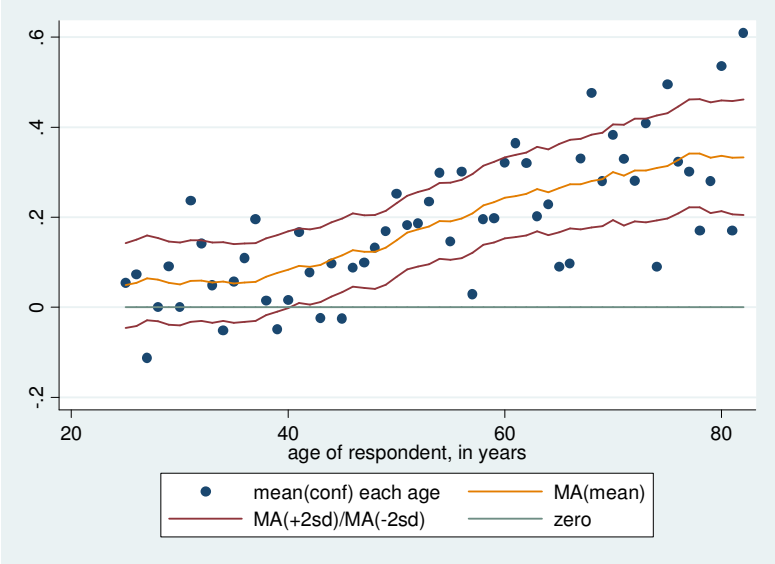


Figure 3: Mean Perceived Health Vulnerability by Age (with moving averages of mean and +/- 2 std.dev.)

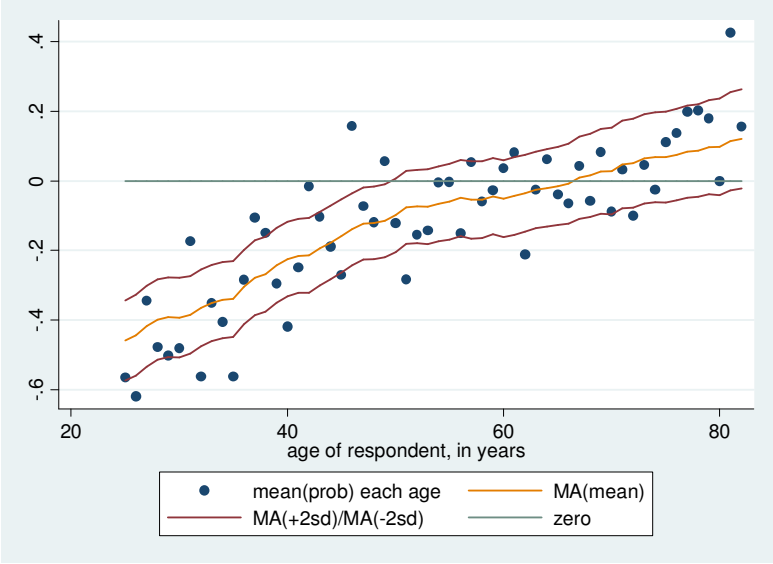


Figure 4: Controllability of cancer

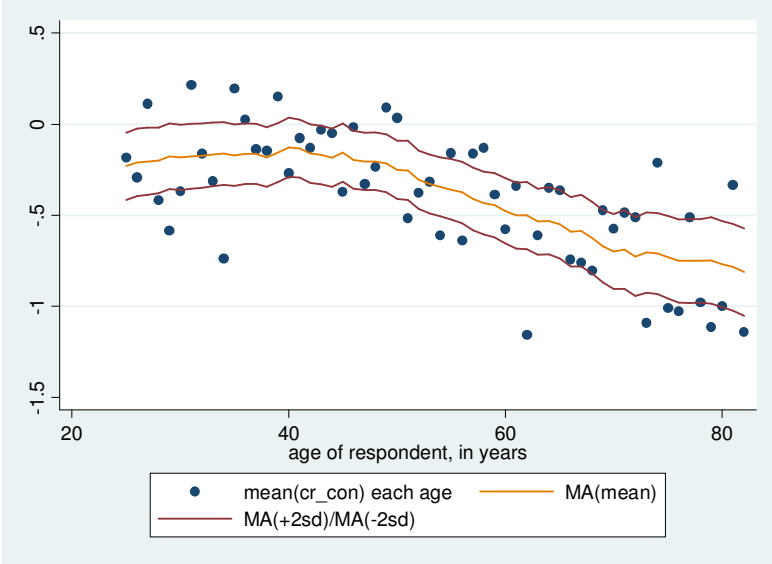
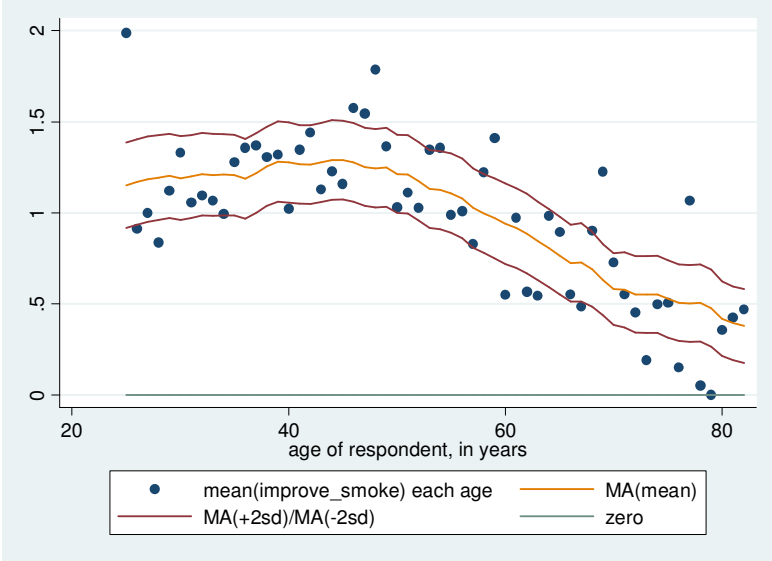


Figure 5: Smokers



NOTE: 47 to 361 individuals in each age group (mean 199)

Figure 6: Own experience with cancer

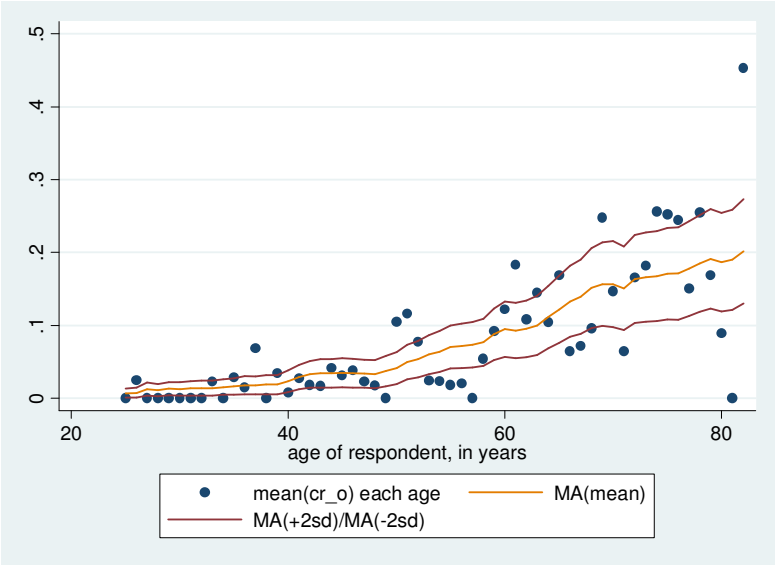


Figure 7: Own experience with diabetes

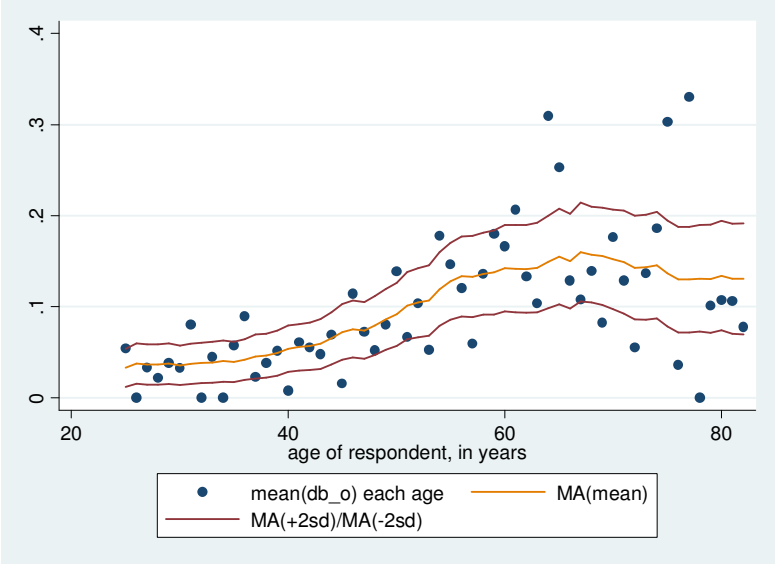


Figure 8: Own or family/friend experience with traffic accidents

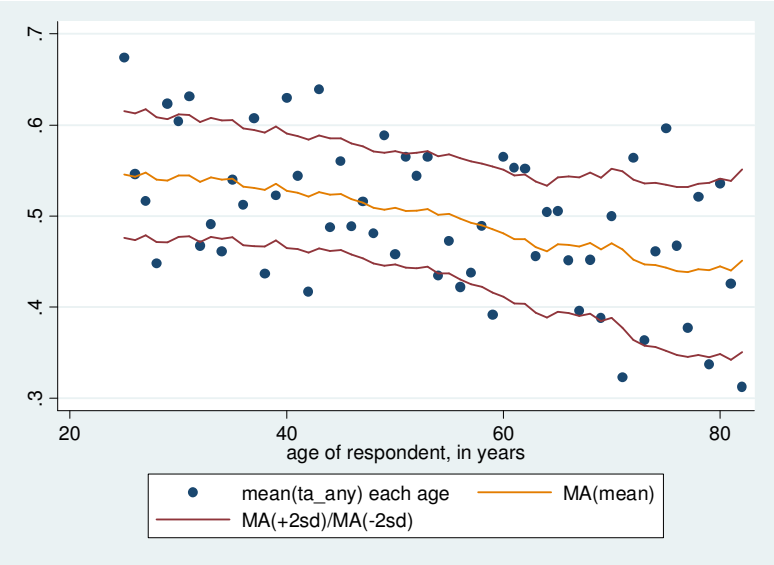


Table 1: Descriptive Statistics

By Illness Profile: [a]	**Included in Illness Profile?**	**Controllability** -2 = very little +2 = a lot	**At Risk** -2 = low risk +2 high risk
Breast Cancer	0.0465	-0.2966	-0.0650
Prostate Cancer	0.0454	”	”
Lung Cancer	0.0912	”	”
Colon Cancer	0.0896	”	”
Skin Cancer	0.0906	”	”
Heart Disease	0.0938	0.2443	0.0057
Heart Attack	0.0933	”	”
Respiratory Disease	0.0891	-0.3881	-0.5481
Stroke	0.0943	-0.0520	-0.3576
Traffic Accident	0.0856	-1.1471	-0.3874
Diabetes	0.0902	-0.2488	-0.4349
Alzheimer's Disease	0.0906	-1.077	-0.9987

By Respondent:	**Mean (Std Dev)**
Age [b] (at time of survey)	50.296 (15.04)
Confidence [b] (in health care)	-.115 (.808)
Vulnerability [b] (in next 20 years)	0.159 (.667)
Could smoke less [c]	.963 (1.619)

[a] Sample Size is 13,696 illness profiles
[b] Sample Size is 1631respondents
[c] Sample Size is 1607 respondents

Table 2: Parsimonious Model

Basic Model Terms:	**Basic Terms**
$(\beta_{00}\times 10^{5})$[linear net income term]	5.369 (7.80)***
$(\beta_{10}\times 10^{9})$[quadratic net income term]	-.198 (3.90)***
(δ_j) disease-specific indicators	(expanded in Table 2a)
$(\alpha_{20})\Delta\Pi_i^{AS}\log(pdvr_i^A+1)$	-17.42 (1.64)
$(\alpha_{10})\Delta\Pi_i^{AS}\left[\log(pdvi_i^A+1)\right]$	(expanded in Table 2b)
$(\alpha_{13})\left[P(sel_i)-\overline{P}\right]\Delta\Pi_i^{AS}\left[\log(pdvi_i^A+1)\right]$	3.517 (2.35)**
$(\alpha_{30})\Delta\Pi_i^{AS}\log(pdvl_i^A+1)$	(expanded in Table 2c)
$(\alpha_{31})age_{i0}\cdot\Delta\Pi_i^{AS}\log(pdvl_i^A+1)$	19.29 (2.73)***
$(\alpha_{32})age_{i0}^2\cdot\Delta\Pi_i^{AS}\log(pdvl_i^A+1)$	-.1614 (2.39)**
$(\alpha_{40})\Delta\Pi_i^{AS}\left[\log(pdvl_i^A+1)\right]^2$	221.5 (2.51)**
$(\alpha_{41})age_{i0}\cdot\Delta\Pi_i^{AS}\left[\log(pdvl_i^A+1)\right]^2$	-8.218 (2.29)**
$(\alpha_{42})age_{i0}^2\cdot\Delta\Pi_i^{AS}\left[\log(pdvl_i^A+1)\right]^2$	.07364 (2.12)**
$(\alpha_{52})age_{i0}^2\cdot\Delta\Pi_i^{AS}\left[\log(pdvi_i^A+1)\right]\cdot\left[\log(pdvl_i^A+1)\right]$	.006295 (2.22)**
Observations	20544
LogL	-10356.924

Table 2a: Parsimonious Model (Alternatives = 20,544)

		Shifters	
	Basic Terms	* Age	* Smoker
Disease Labels: (base case = heart disease)			
heart attack	.6626 (8.40)***	-	-
breast cancer	.653 (5.22)***	-	-
prostate cancer	.6285 (4.66)***	-	-
lung cancer	-.2941 (3.37)***	-	.309 (8.44)***
colon cancer	.2128 (2.91)***	-	-
skin cancer	-.4215 (5.27)***	-	-
respiratory disease	-.4092 (4.63)***	-	.1894 (4.92)***
stroke	.4475 (5.56)***	-	-
traffic accident	-	-.005923 (3.91)***	-
diabetes	.8273 (3.60)***	-.0179 (3.95)***	-
Alzheimer's disease	.4747 (1.44)	-.01119 (1.74)*	-

Table 2b: Parsimonious Model (Alternatives = 20,544)

		Shifters	
Sick Year Terms:	**Basic Terms**	* Age	* Smoker
$(\alpha_{10})\Delta\Pi_i^{AS}\log\left(pdvi_i^A+1\right)$ (base case = heart disease)	-56.65 (4.88)***	-.6759 (2.55)**	-
*heart attack	78.68 (3.24)***	-	-
*breast cancer	-	1.205 (2.10)**	-
*prostate cancer	74.55 (2.19)**	-	-
*stroke	76.24 (3.10)***	-	-

Table 2c: Parsimonious Model (Alternatives = 20,544)

		Shifters	
Lost-life Year Terms:	**Basic Terms**	* Age	* Smoker
$(\alpha_{30})\Delta\Pi_i^{AS}\log\left(pdvl_i^A+1\right)$ (base case = heart disease)	-565.2 (3.22)***	-	-
*Alzheimer's disease	275.5 (2.73)***	-4.964 (2.75)***	-

Table 3: Willingness to pay for a 1 in 1,000,000 risk reduction, by health threat (two scenarios, income = $42,000) [a]						
Profile	**Sudden Death Now**			**10 year latency; sick 5 years, then death**		
Health Threat						
Age now	**30**	**45**	**60**	**30**	**45**	**60**
Heart Disease	8.26 (5.38, 11.65)	8.73 (6.46, 11.35)	7.36 (5.24, 9.96)	9.05 (7.08, 11.32)	7.01 (5.59, 8.62)	4.82 (3.6, 6.26)
Heart Attack	8.22 (5.38, 11.5)	8.68 (6.58, 11.1)	7.29 (5.38, 9.61)	7.58 (5.76, 9.69)	6.09 (4.58, 7.84)	4.4 (3.01, 5.97)
Breast Cancer	7.87 (4.65, 11.34)	8.39 (5.89, 11.25)	6.98 (4.56, 9.57)	8.66 (6.6, 11.04)	6.64 (4.08, 8.35)	4.44 (2.93, 6.03)
Prostate Cancer	7.38 (4.22, 10.74)	7.78 (5.29, 10.58)	6.4 (3.82, 9.08)	7.25 (5.26, 9.42)	5.65 (4.1, 7.44)	4.05 (2.62, 5.69)
Colon Cancer	4.46 (1.76, 7.32)	4.88 (3.09, 6.91)	3.47 (1.72, 5.35)	6.32 (4.67, 8.15)	4.76 (3.6, 6.03)	3.09 (2.03, 4.26)
Lung Cancer	.95 (-2.3, 3.14)	0.97 (-1.05, 2.74)	0.22 (-2.49, 1.29)	2.36 (0.8, 3.99)	0.78 (-0.53, 1.92)	0.04 (-2.34, .3)
…* smoker	10.59 (7.18, 14.26)	11.05 (8.32, 14.21)	9.62 (7, 12.79)	12.43 (9.75, 15.7)	10.93 (8.58, 13.59)	9.24 (7.22, 11.86)
Skin Cancer	0.43 (-3.28, 2.2)	0.36 (-1.98, 1.62)	0.04 (-3.49, 0.26)	1.37 (-0.01, 2.79)	0.16 (-1.55, 0.85)	0 (-3.49, -.67)
Stroke	6.37 (3.45, 9.4)	6.79 (4.85, 8.98)	5.4 (3.6, 7.53)	5.96 (4.17, 7.93)	4.33 (2.99, 5.87)	2.7 (1.29, 4.13)
Respiratory Disease	0.38 (-3.63, 2.12)	0.29 (-2.29, 1.56)	0.03 (-3.74, 0.08)	1.18 (-0.3, 2.81)	0.13 (-1.97, 0.85)	0 (-3.78, -0.69)
…* smoker	5.95 (2.96, 9.13)	6.34 (3.9, 8.99)	5.01 (2.65, 7.48)	7.81 (5.57, 10.51)	6.22 (4.29, 8.34)	4.62 (2.73, 6.72)
Traffic Accident	1.11 (-1.72, 3.48)	0.82 (-1.03, 2.33)	0.07 (-3.2, 0.61)	2.69 (1.42, 4.08)	0.63 (-0.6, 1.64)	0.01 (-3.2, -0.12)
Diabetes	5.3 (2.32, 8.63)	3.36 (1.64, 5.23)	0.3 (-2.35, 1.57)	7.15 (5.1, 9.53)	3.24 (2.1, 4.45)	0.07 (-2.15, 0.47)
Alzheimer's Disease	0.2 (-10.19, 1.6)	0.85 (-3.25, 3.28)	2.24 (-0.5, 4.98)	0.73 (-4.06, 3.35)	0.91 (-1.31, 2.7)	1.03 (-0.55, 2.43)

[a] Cells describe simulated distribution of *WTP* (across 1000 random draws from the joint distribution of the estimated parameters). Since negative values are often an artifact of the functional form and there was no opportunity for any respondent to express a negative *WTP*, we convert negative simulated values to zero before calculating the mean (first entry in each cell). However, we also report the 5th and 95th percentiles of the distribution before negative values are converted to zero.

Appendix I: Entire Parsimonious Model - Table 2 Expanded to Show all Controls

(Alternatives = 20,544)

		Shifters (0=neutral, except for age)						
	Basic Terms	*Confidence	*Vulnerability	* Age	*Vulnerability *Age	*Control	*At Risk	* Smoker
Disease Labels:								
heart disease	-	-.1002 (0.70)	.7906 (3.00)***	-	-.008198 (1.60)	-	-	-
heart attack	.6626 (8.40)***	-	.536 (6.86)***	-	-	-	-	-
breast cancer	.653 (5.22)***	-.3443 (1.95)*	.311 (3.06)***	-	-	-	-	-
prostate cancer	.6285 (4.66)***	-	.3988 (3.71)***	-	-	-	-	-
lung cancer	-.2941 (3.37)***	-	.3465 (4.21)***	-	-	-	-	.309 (8.44)***
colon cancer	.2128 (2.91)***	-	.3377 (4.28)***	-	-	-	-	-
skin cancer	-.4215 (5.27)***	-	.2669 (3.04)***	-	-	-	-	-
respiratory disease	-.4092 (4.63)***	-	.4802 (5.43)***	-	-	-	-	.1894 (4.92)***
stroke	.4475 (5.56)***	-.1737 (1.34)	.538 (6.82)***	-	-	-	-	-
traffic accident	-	-	-	-.005923 (3.91)***	-	-	-	-
diabetes	.8273 (3.60)***	-	.5445 (6.31)***	-.0179 (3.95)***	-	-	-	-
Alzheimer's disease	.4747 (1.44)	-	.2368 (2.82)***	-.01119 (1.74)*	-	-	-	-

Continued on next page…

Appendix 1, continued

		Shifters (0=neutral, except for age)						
Sick Year Terms:	**Basic Terms**	*Confidence	*Vulnerability	* Age	*Vulnerability *Age	*Control	*At Risk	* Smoker
$(\alpha_{10})\Delta\Pi_i^{AS}\log\left(pdvi_i^A+1\right)$	-56.65 (4.88)***	-	-	-	-	-	-	-
*heart disease	-	-35.45 (1.34)	-	-.6759 (2.55)**	-	-	-	-
*heart attack	78.68 (3.24)***	-	-	-	-	-	-	-
*general cancer	-	-	-	-	-	-13.69 (2.72)***		
*breast cancer	-	-	-	1.205 (2.10)**	-	-	-	-
*prostate cancer	74.55 (2.19)**	-	-	-	-	-	-	-
*respiratory disease	-	-	-	-	-	-	12.13 (2.93)***	-
*stroke	76.24 (3.10)***	-	-	-	-	-	-	-
*traffic accident	-	-	-67.29 (2.64)***	-	-	-	-	-
*diabetes	-	-	-	-	-	-	-8.164 (2.04)**	-
*Alzheimer's disease	-	82.32 (2.33)**	-	-	-	8.936 (1.68)*	-	-

Continued on next page…

Appendix 1, continued

		Shifters (0=neutral, except for age)						
Lost-life Year Terms:	**Basic Terms**	*Confidence	*Vulnerability	* Age	*Vulnerability *Age	*Control	*At Risk	* Smoker
$(\alpha_{30})\Delta\Pi_i^{AS}\log\left(pdvl_i^A+1\right)$	-565.2 (3.22)***	-	-	-	-	-	-	-
*any cancer	-	-	-	-	-	-	-10.29 (2.32)**	-
*breast cancer	-	-69.27 (1.75)*	-	-	-	-	-	-
*respiratory disease	-	-	-	-	-	13.47 (3.13)***	-	-
*stroke	-	-53.49 (2.21)**	-	-	-	-13.3 (2.51)**	7.647 (1.59)	-
*diabetes	-	-47.46 (2.61)***	-	-	-	-6.506 (1.54)	-	-
*Alzheimer's disease	275.5 (2.73)***	-90.95 (2.46)**	-	-4.964 (2.75)***	-	-	-	-

Continued on next page…

Appendix 1, continued

Basic Model Terms:	**Basic Terms**	**Shifters (0=neutral, except for age)** *Confidence	*Vulnerability	* Age	*Vulnerability *Age	*Control	*At Risk	* Smoker
$(\beta_{00}\times 10^{5})$[linear net income term]	5.369 (7.80)***	-	-	-	-	-	-	-
$(\beta_{10}\times 10^{9})$[quadratic net income term]	-.198 (3.90)***	-	-	-	-	-	-	-
$(\alpha_{20})\Delta\Pi_i^{AS}\log\left(pdvr_i^A+1\right)$	-17.42 (1.64)	-	-	-	-	-	-	-
$(\alpha_{13})\left[P(sel_i)-\overline{P}\right]\Delta\Pi_i^{AS}\left[\log\left(pdvi_i^A+1\right)\right]$	3.517 (2.35)**	-	-	-	-	-	-	-
$(\alpha_{31})age_{i0}\cdot\Delta\Pi_i^{AS}\log\left(pdvl_i^A+1\right)$	19.29 (2.73)***	-	-	-	-	-	-	-
$(\alpha_{32})age_{i0}^2\cdot\Delta\Pi_i^{AS}\log\left(pdvl_i^A+1\right)$	-.1614 (2.39)**	-	-	-	-	-	-	-
$(\alpha_{40})\Delta\Pi_i^{AS}\left[\log\left(pdvl_i^A+1\right)\right]^2$	221.5 (2.51)**	-	-	-	-	-	-	-
$(\alpha_{41})age_{i0}\cdot\Delta\Pi_i^{AS}\left[\log\left(pdvl_i^A+1\right)\right]^2$	-8.218 (2.29)**	-	-	-	-	-	-	-
$(\alpha_{42})age_{i0}^2\cdot\Delta\Pi_i^{AS}\left[\log\left(pdvl_i^A+1\right)\right]^2$	.07364 (2.12)**	-	-	-	-	-	-	-
$(\alpha_{52})age_{i0}^2\cdot\Delta\Pi_i^{AS}\left[\log\left(pdvi_i^A+1\right)\right]\cdot\left[\log\left(pdvl_i^A+1\right)\right]$	.006295 (2.22)**	-	-	-	-	-	-	-
Observations	20544							
LogL	-10356.924							

Absolute value of z statistics in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

Appendix II: Model Progression and Results

We develop our final working specification through a process whereby we start with a heavily parameterized model with a wide array of potentially relevant interaction terms. Given that intuition suggests that the type of illness should have some bearing on the (dis)utility from a prospective sick-year, we first allow the coefficients on the sick-year terms to vary. Results for all of our preliminary models have been relegated to Appendix II. Here, we will merely review some of the more notable results.

Model 1 (for which the key parameters of interest are displayed in the first column of Appendix Table A.1) shows the consequences of allowing the simple coefficient on the linear term in the log of discounted sick-years to vary systematically according to the disease label assigned to the illness profile. We use heart disease as the base case and all other marginal utilities are differentials relative to the (dis)utility from a discounted sick-year with heart disease. The estimated coefficients suggest that a year with breast cancer (for female respondents) is viewed as contributing greater disutility than a year with heart disease. The (dis)utilities per sick year appear to be much less for all other disease labels, and the positive differentials are sufficiently large for skin cancer and traffic accidents that the result imply positive utilities per discounted sick year for these illnesses, which is implausible. So we quickly begin to search for necessary controls.

Given the implausible effects on the marginal (dis)utility of sick-years for some illness labels, we explore (in Model 2, occupying the rest of Table A.1) whether they may be heterogeneity in illness effects according to a selection of relevant respondent attributes or attitudes. The results for Model 2 span seven columns, since we display the coefficients on the additional interaction terms lined up with the disease labels.

Our empirical results suggest that the marginal (dis)utility of a sick-year for many diseases is statistically indistinguishable from that associated with heart disease. These results suggest that colon cancer, skin cancer, respiratory disease, and Alzheimer's disease have statistically significantly different baseline parameters than heart disease, but there appears to be additional heterogeneity with age for breast cancer (a declining concern) and diabetes (also declining). The disutility of sick-years from heart disease also appears to be lower when the respondent is older.

In Table A.2, we introduce the dummy variables for disease labels not as shifters on the marginal (dis)utility of a sick-year, but as sources of a "lump" of indirect utility associated with the program in question. In other work, we have determined that a dummy variable simply for "any program other than the status quo" has a statistically significant effect on the respondent's propensity to choose among the different options. People appear to prefer to derive utility from our proposed risk-reduction options that is distinct from the attributes of the program (cost, risk reduction) and the nature of the health risk to be reduced (the illness profile in question). Instead of using a single unspecified dummy variable for "any program," we substitute a set of indicator variables for each illness label. All indicators are set to zero for the status quo "no program" option in each choice set.

The results in Table A.2 pertain to a single specification, arranged so that the variables corresponding to each disease appear in the same row. This model exhibits a somewhat greater number of statistically significant effects than were apparent in Table A.1. Half of the other disease labels bear coefficients which indicate that the effect of the label is statistically significantly different than that for heart disease. Each significant disease label contributes a positive amount to utility. Table A.2 reveals that smoking status has a large and very significant

effect on preferences for risk-reduction programs pertaining to lung cancer and respiratory disease.

Older respondents seem to attach statistically significantly higher utility to programs which reduce the risk of heart disease, but lower utility to programs to reduce the risk of breast cancer, traffic accidents, and diabetes. For a respondent with "neutral" attitudes who is a non-smoker, all of the other interaction terms will be zero. Also, despite the introduction of these rich opportunities to capture heterogeneity, the coefficients on the basic model remain for the most part, stable and statistically significant.

Having considered two alternative ways to incorporate illness labels into our model, we now include both methods, in the same model, to assess whether one strategy seems to dominate. Table A.3 displays the key results for this more-highly parameterized model with both types of shifters. It appears that statistically significant disease label effects may be more prevalent among the "lumps" of utility associated with each risk-reduction program than they are among the terms which involve shifts in the marginal utility of log discounted sick-years, but there are many individually statistically significant coefficients in each portion of the model.

Adding illness labels as factors which contribute lumps of utility (as in Table A.1) to the model in Table A.2 increases the log likelihood by roughly 254 points (yielding a chi-squared test statistic of over 500, which vastly exceed the relevant critical value for 50 degrees of freedom (which is less than 80). Similarly, if we modify the sample so that the same 20,544 alternatives are included for each model, the log-likelihood increases by 314 points (yielding a chi-squared test statistic of over 600). This strongly rejects the null hypothesis that we do not need to allow for heterogeneity by illness label in the marginal utility of log discounted sick-years.

Having considered the possibilities that heterogeneity in demand by illness label may affect the marginal (dis)utility of discounted sick-years, and/or the basic indirect utility associated with each program *regardless* of the time profile of the illness, we expand the model, finally, to allow for heterogeneity by illness label in the marginal (dis)utility of discounted lost life-years. Results are shown in Table A.4. Again, numerous differences by illness label are seen in all three types of generalizations.

The overwhelming number of coefficients in the most extensive model needs to be reduced. Given the randomized design of the program attributes, there is much less risk that multicollinearity will complicate the process of identifying stable and statistically significant illness label effects by pruning away persistently insignificant variables. Of course, the variables which we use to control further, for respondent attitudes, are correlated to a certain extent, but we will be zeroing out the influence of these variables in our simulations. These are merely control variables, included where necessary to help us better discern the tendencies in the data which can be identified for a respondent with "neutral" attitudes on all dimensions, for selected specific age levels.

Table A.1: Sick Year Shifters Only (Alternatives = 20,544)

Sick Year Terms:	**Model 1**	**Model 2**	*Confidence	*Vulnerability	*Age	*Vulnerability *Age	*At Risk	*Control
$(\alpha_{10})\Delta\Pi_i^{jS}\log\left(pdvi_i^j+1\right)$	-95.35	-191.7	-	-	-	-	-	-
	(7.45)***	(3.84)***						
* heart disease	-	-	-21.51	-151.4	2.077	2.155	4.237	-.488
			-1.26	(2.88)***	(2.24)**	(2.12)**	-0.7	(0.07)
* heart attack	14.22	129.8	10.82	-132.9	-0.3226	0.828	-	-
	-0.66	-1.49	-0.36	-1.42	-0.2	-0.46		
* any cancer	-	-	-	-	-	-	-0.6119	-7.972
							-0.12	(1.29)
* breast cancer	-13.77	-137	-3.815	-89.64	4.236	1.085	-	-
	-0.6	-1.47	-0.12	-0.93	(2.62)***	-0.61		
* prostate cancer	1.983	123.3	21.88	-145.1	-0.504	2.492	-	-
	-0.09	-1.37	-0.68	-1.64	-0.31	-1.45		
* lung cancer	40.38	99.57	21.97	-34.34	0.9399	-0.3927	-	-
	(2.56)**	-1.61	-1.16	-0.59	-0.9	-0.34		
* colon cancer	21.72	150.6	35.38	-65.12	-0.5888	0.647	-	-
	-1.42	(2.49)**	(2.01)**	-1.15	-0.61	-0.6		
* skin cancer	110.1	191.6	-13.1	20.3	0.6026	-0.9654	-	-
	(6.56)***	(2.93)***	-0.65	-0.32	-0.55	-0.77		
* respiratory disease	97.75	241.2	35.56	-78.59	-0.6714	0.09412	11.85	-.7585
	(5.49)***	(3.30)***	-1.59	-0.99	-0.55	-0.06	(2.44)**	(0.14)
* stroke	45.68	46.7	17.16	-188.4	1.747	1.756	5.56	-11.3
	(2.12)**	-0.5	-0.57	(1.76)*	-1.06	-0.88	-0.87	(1.62)
* traffic accidents	107.9	99.05	-13.63	51.94	2.679	-2.395	5.291	-6.988
	(4.80)***	-1.13	-0.42	-0.57	-1.6	-1.22	-0.98	(1.27)
* diabetes	67.22	61.45	-49.46	-101.1	2.345	0.5836	-8.056	-3.786
	(3.77)***	-0.88	(2.14)**	-1.4	(1.91)*	-0.43	(1.70)*	(0.70)
* Alzheimer's disease	76.14	312.1	17.39	31.44	-2.241	-0.927	-1.943	15.01
	(4.20)***	(3.99)***	-0.76	-0.39	(1.76)*	-0.65	-0.34	(2.40)**
Observations	20544	20544						
LogL	-10610.6	-10537.1						

Additional parameters are included in the model and retain the same signs and significance levels as they display in Table 2, the model without heterogeneity by disease label: Linear and quadratic income, recovered year, selection correction, interactions between age and interaction between sick years and lost life years, and higher order lost-life year and interaction terms.
Absolute value of z statistics in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

Table A.2: Simple Disease Label Effects Only (Alternatives = 20,544)

		Shifters (0=neutral, except for age)				
Disease Labels:	Basic Terms	*Confidence	*Vulnerability	*Age	*Vulnerability *Age	*Smoke
heart disease	-	-.0451	1.059	.01016	-.01329	-
		(0.49)	(3.88)***	(7.36)***	(2.50)**	
heart attack	.9605	.1938	.8664	-.006974	-.005922	-
	(4.00)***	(2.07)**	(3.09)***	(1.49)	(1.07)	
breast cancer	1.463	-.08836	.5559	-.01756	-.003789	-
	(4.72)***	(0.71)	(1.53)	(2.93)***	(0.53)	
prostate cancer	.6704	-.06969	.9738	-.002084	-.01167	-
	(2.14)**	(0.54)	(2.65)***	(0.34)	(1.58)	
lung cancer	.1108	-.1194	.3749	-.004723	.0004623	.2952
	(0.43)	(1.21)	(1.31)	(0.98)	(0.08)	(7.91)***
colon cancer	.4528	-.007631	.4561	-.001773	-.002096	-
	(1.94)*	(0.08)	(1.64)	(0.40)	(0.38)	
skin cancer	.04743	-.008459	.3777	-.006317	-.00153	-
	(0.18)	(0.08)	(1.23)	(1.26)	(0.25)	
respiratory disease	-.04607	-.0806	.6294	-.004015	-.002302	.1804
	(0.17)	(0.78)	(2.02)**	(0.81)	(0.37)	(4.60)***
stroke	.731	.03887	.9462	-.00574	-.007716	-
	(3.09)***	(0.42)	(3.43)***	(1.25)	(1.42)	
traffic accident	.3349	.08645	.1857	-.01069	-.0004559	-
	(1.23)	(0.81)	(0.61)	(2.04)**	(0.07)	
diabetes	1.159	.2556	1.039	-.0219	-.009477	-
	(4.65)***	(2.58)***	(3.45)***	(4.42)***	(1.57)	
Alzheimer's disease	-.0148	-.0333	.2189	.0003553	.0006567	-
	(0.06)	(0.34)	(0.76)	(0.08)	(0.12)	
Observations	20544					
LogL	-10370.799					

Additional parameters are included in the model and retain the same signs and significance levels as they display in Table 2, the model without heterogeneity by disease label: Linear and quadratic income, sick years, recovered years, lost-life years, selection correction, interactions between age and sick years, interaction between sick years and lost life years, and higher order lost-life year and interaction terms. Absolute value of z statistics in parentheses ; * significant at 10%; ** significant at 5%; *** significant at 1%

Table A.3: Simple Disease Label Effects *and* Sick Year Shifters (Alternatives = 20,544)

		Shifters (0=neutral, except for age)						
Disease Labels:	Basic Terms	*Confidence	* Vulnerability	*Age	* Vulnerability *Age	*Smoke	*At Risk	*Control
heart disease	-	-.3143	1.056	.01409	-.01321	-	-	-
		(2.17)**	(2.49)**	(7.07)***	(1.54)			
heart attack	1.385	.2832	.8113	-.01311	-.005392	-	-	-
	(4.90)***	(2.65)***	(2.49)**	(2.36)**	(0.84)			
breast cancer	1.645	-.3666	.2653	-.01894	.003763	-	-	-
	(3.31)***	(1.80)*	(0.44)	(1.94)*	(0.30)			
prostate cancer	1.201	-.05958	.8127	-.009229	-.004269	-	-	-
	(2.47)**	(0.29)	(1.38)	(0.96)	(0.35)			
lung cancer	.4581	-.2087	.564	-.01338	-.004338	.298	-	-
	(1.20)	(1.35)	(1.28)	(1.80)*	(0.48)	(7.91)***		
colon cancer	.8928	.3423	.02574	-.01174	.006426	-	-	-
	(2.54)**	(2.30)**	(0.06)	(1.71)*	(0.73)			
skin cancer	.2758	-.1693	.6447	-.01257	-.008539	-	-	-
	(0.69)	(0.98)	(1.38)	(1.58)	(0.90)			
respiratory disease	.7035	.01013	1.189	-.01815	-.01191	.1833	-	-
	(1.81)*	(0.06)	(2.44)**	(2.41)**	(1.23)	(4.63)***		
stroke	.9261	.07068	.9136	-.007343	-.007812	-	-	-
	(3.33)***	(0.66)	(2.89)***	(1.35)	(1.25)			
traffic accident	.5665	.06044	.3002	-.01432	-.005281	-	-	-
	(1.73)*	(0.48)	(0.85)	(2.27)**	(0.75)			
diabetes	2.159	.2876	1.783	-.04211	-.02311	-	-	-
	(5.36)***	(1.67)*	(3.40)***	(5.19)***	(2.12)**			
Alzheimer's disease	.7238	.05709	.5165	-.01478	-.004757	-	-	-
	(1.93)*	(0.36)	(1.13)	(2.00)**	(0.52)			

Continued on next page…

Table A.3 continued

Sick Year Terms:	Basic Terms	Shifters (0=neutral, except for age) *Confidence	* Vulnerability	*Age	* Vulnerability *Age	*Smoke	*At Risk	*Control
$(\alpha_{10})\Delta\Pi_i^{AS}\log(pdvi_i^A+1)$	-286.6 (5.57)***	-	-	-	-	-	-	-
* heart disease	-	-72.6 (2.63)***	-46.83 (0.56)	5.531 (5.47)***	.5728 (0.34)	-	3.836 (0.63)	.2977 (0.04)
* heart attack	361 (3.54)***	64.32 (1.84)*	-72 (0.68)	-1.117 (0.60)	.8381 (0.40)	-	-	-
* any cancer	-	-	-	-	-	-	-1.176 (0.23)	-8.361 (1.35)
* breast cancer	238.7 (1.69)*	-87.15 (1.71)*	-102 (0.64)	.8789 (0.35)	2.425 (0.80)	-	-	-
* prostate cancer	368.4 (2.73)***	12.41 (0.24)	-51.18 (0.36)	-1.359 (0.54)	2.157 (0.75)	-	-	-
* lung cancer	273.8 (3.12)***	-19.55 (0.64)	28.46 (0.31)	-.91 (0.58)	-.8605 (0.47)	-	-	-
* colon cancer	309.5 (3.67)***	87.15 (3.04)***	-125.8 (1.38)	-1.501 (1.05)	2.346 (1.32)	-	-	-
* skin cancer	251.5 (2.77)***	-34.6 (1.06)	49.71 (0.53)	-.4511 (0.28)	-1.49 (0.78)	-	-	-
* respiratory disease	409 (3.91)***	30.79 (0.89)	157.7 (1.26)	-2.884 (1.61)	-2.66 (1.18)	-	12.83 (2.64)***	4.302 (0.78)
* stroke	212.5 (1.96)*	21.86 (0.62)	-75.68 (0.63)	1.715 (0.90)	.781 (0.34)	-	4.883 (0.77)	-11.14 (1.60)
* traffic accidents	214.7 (2.10)**	-11.23 (0.30)	43.9 (0.41)	1.154 (0.59)	-2.661 (1.18)	-	3.557 (0.66)	-8.583 (1.56)
* diabetes	457.9 (4.42)***	10.34 (0.26)	194 (1.59)	-4.148 (2.27)**	-3.517 (1.49)	-	-7.253 (1.52)	-6.465 (1.19)
* Alzheimer's disease	394.3 (3.46)***	25.77 (0.69)	69.95 (0.54)	-2.912 (1.53)	-1.257 (0.54)	-	-2.55 (0.45)	14.62 (2.33)**
Observations	20544							
LogL	-10296.213							

Additional parameters are included in the model and retain the same signs and significance levels as they display in Table 2: Linear and quadratic income, recovered years, lost-life years, selection correction, interactions between age and sick years, interaction between sick years and lost life years, and higher order lost-life year and interaction terms. Absolute value of z statistics in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

Table A.4: Full Model (Simple Disease Label Effects, Sick Year and Lost Life-year Shifters, Alternatives = 20,544)

		Shifters (0=neutral, except for age)						
	Basic Terms	*Confidence	*Vulnerability	*Age	*Vulnerability *Age	*Smoker	*Control	*At Risk
Disease Labels:								
heart disease	-	-.2628	1.278	.01453	-.01906	-	-	-
		(1.70)*	(2.82)***	(6.91)***	(2.09)**			
heart attack	1.182	.1753	.7076	-.007819	-.003539	-	-	-
	(3.24)***	(1.10)	(1.55)	(1.08)	(0.39)			
breast cancer	1.133	-.467	.3414	-.0102	.00266	-	-	-
	(2.08)**	(2.10)**	(0.52)	(0.96)	(0.20)			
prostate cancer	1.242	-.192	.8572	-.009694	-.00599	-	-	-
	(2.44)**	(0.88)	(1.38)	(0.96)	(0.46)			
lung cancer	.5077	-.21	.5299	-.01453	-.002589	.3014	-	-
	(1.22)	(1.23)	(1.10)	(1.81)*	(0.26)	(7.95)***		
colon cancer	.9681	.3181	.03662	-.01346	.006045	-	-	-
	(2.54)**	(2.01)**	(0.08)	(1.81)*	(0.62)			
skin cancer	.4673	-.1569	.6625	-.01696	-.006844	-	-	-
	(1.07)	(0.84)	(1.29)	(1.97)**	(0.65)			
respiratory disease	.7097	.04212	1.276	-.01883	-.01343	.1821	-	-
	(1.72)*	(0.25)	(2.37)**	(2.34)**	(1.25)	(4.59)***		
stroke	.8022	-.1859	1.263	-.004025	-.01792	-	-	-
	(2.23)**	(1.21)	(2.76)***	(0.57)	(1.95)*			
traffic accident	.6699	-.06071	-.087	-.01419	.003783	-	-	-
	(1.59)	(0.34)	(0.17)	(1.71)*	(0.36)			
diabetes	2.293	.224	2.144	-.04394	-.0306	-	-	-
	(5.21)***	(1.23)	(3.69)***	(4.90)***	(2.52)**			
Alzheimer's disease	1.025	-.08928	.6156	-.02042	-.006368	-	-	-
	(2.58)***	(0.52)	(1.24)	(2.59)***	(0.64)			

Note: Additional parameters are included in the model and retain the same signs and significance levels as they display in Table 2: Linear and quadratic income, recovered years, selection correction, interactions between age and sick years, the interaction between sick years and lost life years, and higher order lost-life year and interaction terms. Absolute value of z statistics in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

Continued on next page…

Table A.4, continued

Sick Year Terms:	Basic Terms	**Shifters (0=neutral, except for age)** *Confidence	* Vulnerability	*Age	* Vulnerability *Age	*Smoker	*Control	*At Risk
$(\alpha_{10})\Delta\Pi_i^{AS}\log(pdvi_i^A+1)$	-273.2 (4.39)***	-	-	-	-	-	-	-
*heart disease	-	-75.95 (2.60)***	-88.64 (0.99)	4.997 (4.10)***	1.639 (0.90)	-	2.449 (0.33)	5.769 (0.87)
*heart attack	341 (3.13)***	59.19 (1.66)*	-75.01 (0.70)	-1.178 (0.62)	.7862 (0.37)	-	-	-
*any cancer	-	-	-	-	-	-	-8.089 (1.20)	3.37 (0.59)
*breast cancer	380 (2.39)**	-72.95 (1.31)	-93.52 (0.55)	-1.863 (0.69)	2.392 (0.75)	-	-	-
*prostate cancer	352.2 (2.17)**	53.26 (0.85)	-100.9 (0.57)	-1.497 (0.51)	3.315 (0.98)	-	-	-
*lung cancer	250.3 (2.54)**	-21.54 (0.68)	38.22 (0.39)	-.8414 (0.51)	-1.26 (0.64)	-	-	-
*colon cancer	280.9 (2.98)***	92.62 (3.03)***	-136.1 (1.46)	-1.291 (0.86)	2.616 (1.42)	-	-	-
*skin cancer	202.5 (1.99)**	-37.39 (1.08)	54.17 (0.54)	.1566 (0.09)	-2.01 (0.99)	-	-	-
*respiratory disease	412.1 (3.51)***	23.98 (0.64)	144.5 (1.10)	-3.297 (1.72)*	-2.528 (1.06)	-	-2.359 (0.40)	14.61 (2.77)***
*stroke	185.4 (1.60)	12.53 (0.35)	-54.33 (0.44)	1.817 (0.93)	.2489 (0.11)	-	-6.593 (0.87)	1.023 (0.15)
*traffic accident	208.8 (1.90)*	-15.61 (0.41)	22.91 (0.21)	.8777 (0.44)	-2.232 (0.97)	-	-7.99 (1.33)	6.511 (1.11)
*diabetes	407.7 (3.42)***	38.56 (0.82)	124.5 (0.94)	-4.163 (2.11)**	-2.274 (0.91)	-	-3 (0.51)	-8.932 (1.72)*
*Alzheimer's disease	273.2 (2.12)**	73.93 (1.77)*	11.61 (0.08)	-1.243 (0.60)	-.5095 (0.20)	-	13.17 (1.93)*	-5.859 (0.95)

Continued on next page…

Table A.4, continued

Lost Life-Year Terms:		**Shifters (0=neutral, except for age)**						
		*Confidence	* Vulnerability	*Age	* Vulnerability *Age	*Smoker	*Control	*At Risk
$(\alpha_{30})\Delta\Pi_i^{AS}\log\left(pdvl_i^A+1\right)$	-588.2 (2.70)***	-	-	-	-	-	-	-
*heart disease	-	26.59 (0.70)	177.3 (1.55)	1.758 (0.61)	-4.529 (1.90)*	-	-3.476 (0.47)	-5.42 (0.87)
*heart attack	-5.615 (0.05)	-21.87 (0.84)	-23.12 (0.29)	1.826 (0.69)	.5835 (0.36)	-	-	-
*any cancer	-	-	-	-	-	-	-2.17 (0.35)	-11.07 (2.09)**
*breast cancer	-245.9 (1.70)*	-42.62 (0.94)	24.63 (0.18)	5.492 (1.76)*	-.2938 (0.11)	-	-	-
*prostate cancer	72.63 (0.52)	-80.97 (1.67)*	78.29 (0.61)		-2.013 (0.75)	-	-	-
*lung cancer	83.94 (0.70)	-.1182 (0.00)	-10.39 (0.10)	-.3346 (0.12)	.8254 (0.40)	-	-	-
*colon cancer	106.5 (0.88)	-14.24 (0.43)	33.45 (0.33)	-.9154 (0.33)	-.7189 (0.34)	-	-	-
*skin cancer	165.7 (1.33)	4.98 (0.16)	7.409 (0.07)	-2.116 (0.76)	1.044 (0.45)	-	-	-
*respiratory disease	27.96 (0.21)	23.01 (0.57)	52.7 (0.40)	.5354 (0.18)	-.6867 (0.26)	-	15.67 (2.85)***	-3.674 (0.75)
*stroke	8.59 (0.07)	-62.23 (2.29)**	98.15 (1.10)	1.564 (0.58)	-2.561 (1.46)	-	-12.8 (1.77)*	8.593 (1.32)
*traffic accident	66.42 (0.57)	-28.86 (0.98)	-82.59 (0.99)	.6184 (0.23)	2.181 (1.22)	-	-3.239 (0.57)	-7.596 (1.38)
*diabetes	116.9 (0.97)	-41.31 (1.20)	155.9 (1.57)	-.2036 (0.07)	-2.992 (1.48)	-	-7.497 (1.35)	4.329 (0.88)
*Alzheimer's disease	285.1 (2.00)**	-101.7 (2.50)**	116.3 (0.87)	-3.743 (1.24)	-1.677 (0.69)	-	6.899 (1.08)	7.343 (1.23)
Observations	20544							
LogL	-10252.933							