Version: 7/23/2012

Handbook to Accompany "Demand for Health Risk Reductions" and Related Papers

By TRUDY ANN CAMERON AND J.R. DESHAZO*

Abstract

This document contains supplementary materials to accompany all papers using the "private choices" survey associated with our larger project on the valuation of morbidity and mortality risk reductions for policy analysis. Page limits for journal articles do not generally permit the amount of detail that is necessary to assure readers about the quality of the underlying data or to explain the reasoning behind such important issues as survey and experimental design decisions, exclusion criteria, different modeling assumptions, alternative specifications, sample selection corrections, and sensitivity/robustness assessments. We also provide a full disclosure of all substantive peer review comments received on the main paper from this project, along with our detailed responses to these questions or criticisms, whether or not these responses were solicited. Such disclosure, while unusual, assures transparency of the peer review process. It is intended to convey the thoroughness of the review process for potential consumers of our results and to provide insights for other researchers who may contemplate future studies in this area.

This is a "living document" in that we plan to revise/update/expand its contents as the related research papers move through the review process, so be sure to check the automatic date on the document before comparing versions. To guide the reader, we include a detailed table of contents, lists of tables and figures, and a comprehensive index.

JEL Classifications: JEL I1, H43, Q51

* Cameron: Department of Economics, University of Oregon, Eugene, OR 97403-1285 (e-mail: cameron@uoregon.edu; DeShazo: Department of Public Policy, School of Public Affairs, 3250 Public Policy Building, UCLA, Los Angeles, CA 90095-1656 (e-mail: deshazo@ucla.edu). We thank Vic Adamowicz, Richard Carson, Maureen Cropper, Baruch Fischhoff, Jim Hammitt, F. Reed Johnson, Alan Krupnick, V. Kerry Smith, and Kip Viscusi, as well as numerous conference and seminar participants, and several discussants and anonymous referees for helpful comments on this paper and other papers which use these data. Rick Li implemented our survey very capably. Ryan Bosworth, Graham Crawford, Dan Burghart and Ian McConnaha have provided able assistance. This research has been supported by the US Environmental Protection Agency (R829485), Health Canada (H5431-010041/001/SS), the National Science Foundation (SES-0551009), and the Mikesell Foundation at the University of Oregon. This work has not been formally reviewed by any of these entities. Any remaining errors are our own.

Handbook to Accompany
"Demand for Health Risk Reductions" and Related Papers

Contents

List of Tables

NOTE: for updates, delete old table. Go to “References,” “Insert table of figures”; set caption label to “Table”; check “include label and number”; do NOT replace existing Table of Figures

List of Figures

NOTE: for updates, delete old list. Go to References/Insert Table of Figures; Use Caption label: Figure; Select Options/Style and choose Heading 6; do NOT replace existing Table of Figures

This document supports the full range of research papers produced using our U.S. "private choices" survey. An inventory of these papers is contained in Section 6. This survey was one of four health-related surveys conducted with external research support from the US Environmental Protection Agency (R829485) and Health Canada (Contract H5431-010041/001/SS), with continuation of the research supported by a grant from the National Science Foundation (SES-0551009). From its inception, our work on this project to this point has spanned almost a decade, so there is far too much material to include in any single journal-length paper. Some of this material documents auxiliary analyses to support parenthetical or footnote material in the various papers derived from this survey. In other cases, the material was generated in response to queries from referees of the various manuscripts as they have moved through the review process. We gratefully acknowledge the concerns and suggestions of our various referees on our different papers, but in some cases (where our additional analyses proved to make little difference to the paper in question), we have elected to report the additional results in detail only in this document. In other cases, the concerns of a referee have actually been irrelevant to our study, so we have likewise produced expanded explanations as to why this is the case, since similar misconceptions may arise in the context of other papers employing these data.

1 Survey Design & Development

Equation Section (Next)

In this section, we provide a succinct overview, for the casual reader, of the survey development process and describe the final survey instrument that we employ. Many of these brief points are pursued in greater detail in subsequent sections. In Section 1.1, we describe the underlying goals and guiding principles for our survey, the cognitive interviews, peer review of survey instrument, and the pre-testing that preceded the fielding of the final instrument. We also discuss some issues related to constraints on the allowable duration for our survey (i.e. panelist minutes) and respondent cognitive constraints, as well as some considerations concerning the survey sample.

In Section 1.2, we discuss the configuration of the survey instrument, which is structured around four modules: (1) risk perceptions and risk-related behaviors, (2) a tutorial for risk changes and illness profiles, (3) the presentation of the choice sets, and (4) a debriefing and follow-up module. Throughout, we discuss design issues and potential biases that we explicitly sought to address when designing the survey. Finally, in Section 4 and 5, we discuss the respondents' health profile survey and the socio-economic profile survey respectively.

A brief statement of our broader research objectives may assist the reader in interpreting the structure and format of our survey instrument. Forward-looking individuals face a portfolio of distinct health risks over their life-time. In each year of their life, their probability associated with each illness or injury changes as does their probability of experiencing a particular health state. Individuals may avail themselves of a wide range of public policies and privately-available medical and behavioral programs that reduce specific types of risks. The vast majority of these policies and programs change the probability that an individual will experience a particular illness (or suite of illnesses) by changing the probability of a particular time profile of health states over their lifespan (see Picone et al. (2004)). For example, choosing to participate in regular prostrate exams or mammogram programs changes these individuals' expected time profiles of the health states associated with these illnesses.

In contrast, however, traditional mortality valuation studies (such as a hedonic wage[1] and recent stated preference studies[2]) do not collect data on the most common choice dynamics which involve substituting across multiple types of risk while allocating risk reductions across time periods. Rather, traditional studies collect data on choices regarding a single risk reduction for the current period only. As such, these studies are unable to model individuals making choices that substitute across distinct types of risks (see Dow et al. (1999)). They are also unable to observe individuals making choices that change their inter-temporal allocation of health risks across future years of their remaining lifetime (Hamermesh (1985)).

In light of this large gap in the literature, our overarching goal was to design a survey that observed individuals' choices over multiple sources of distinct risks. Our survey also seeks to observe individuals' choices with respect to options that change their probability of experiencing future undesirable health states over different periods of time. This is important because most mortality-reduction policies and programs do not "save" lives; rather, they extend life by deferring the future onset of illnesses that result in morbidity and premature mortality.

In this survey we present respondents with an illness-specific health-risk reduction program that involves diagnostic screening, remedial medications, and lifestyle changes that would reduce their probability of experiencing that illness profile. Individuals must pay an annual fee to participate in each risk-reducing program. They are asked to choose between one of two risk-reducing programs (each associated with a different illness profile) or to reject both programs. An advantage of this choice setting is that the individual faces a portfolio of health risks that resemble those they actually face. Through their choices, individuals reveal trade-offs across specific illnesses and a full continuum of health states of different durations. We also observe them strategically allocating expenditures for risk-mitigating programs across the current year and future years of their remaining life (Ehrlich and Chuma (1990); Ehrlich (2000)). Individuals' fundamental object of choice is the probability of spending a year in various health states. By using stated preference methods to gain a window upon these previously unobservable types of intertemporal choices, we are able to estimate the marginal value of a sick year and lost life year in a carefully controlled setting.

A second goal of the survey was to generate choice data that could be used to characterize the full continuum of health state outcomes over time associated with typical public policies. Individuals' observed choices permit us to evaluate infinite combinations of illness profiles, including for example, (1) a period of shorter-term morbidity followed by recovery, (2) a period of longer-term morbidity followed by recovery, (3) a combination of shorter-term morbidity followed by premature mortality, (4) a combination of longer-term morbidity followed by premature mortality, and (5) immediate mortality. With the estimates of this continuum of values for statistical illness profiles, this survey design permits us to more accurately value the actual benefits of different types of public policies to improve environmental, health, and safety outcomes.

A third goal in our survey design was to evaluate the comprehensive effects of a wide variety of different sources of heterogeneity on demand for risk reductions across individuals. Sources of heterogeneity may include individuals' age, health status, discount rate, incomes, *ex*

---

[1] For a review of revealed preference studies see Viscusi (1993); Mrozek and Taylor (2002); de Blaeij et al. (2003) and Viscusi and Aldy (2003).

[2] Recent stated preference studies include Sloan et al. (1998); Johnson et al. (2000); Krupnick et al. (2002); Chestnut et al. (2003); Hammitt and Liu (2004). We studied the survey instruments used in each of these studies carefully when preparing this survey instrument.

ante defensive, averting and mitigating behaviors, and their *ex ante* information on illness specific risks and their subject illness time profiles for specific illnesses. For a sample of studies that explore some subsets of these sources of heterogeneity see Shogren and Crocker (1999); Quiggin (2002); Viscusi (2003); Aldy and Viscusi (2003); Smith et al. (2004); and Viscusi and Aldy (2003).

A fourth goal was to generate a data set that was more representative of the US population than those used in past revealed and stated preferences studies. Many studies are based on non-representative sub-populations (e.g., working age men or convenience samples) while our sample is of the general population aged 25 and older, including men and women, as well as a wide range of ethnicities, and income groups. In addition, most studies focus upon only one source of health risk (typically, accidental on-the-job death).[3]

Finally, a fifth goal was a survey design that accommodates the widest array, to date, of robustness and validity checks as well as sensitivity analysis for a risk valuation study to date. Such checks include assessing risk comprehension, scope effects, order effects, scenario rejection, and sample selection biases. Through our survey design, we also endeavor to mitigate hypothetical bias associated with incentive incompatibility and bias associated with omitting relevant substitute risks and future health states.

1.1 Survey development

We faced several challenging tasks and constraints as we developed this survey instrument. We needed a way of describing the probabilistic time profiles of health states associated with different health risks. These probabilistic profiles would have to be framed within individuals' remaining expected lifespan. We then needed to identify a program that credibly reduced the risk of a wide range of health risks and for which there was a generally acceptable payment vehicle. Of course, we also faced the task of communicating changes in the risk levels associated with each program. In light of these challenges, we developed the initial version of the survey only after an extensive review of the existing literature to March 2002.

1.1.1 Goals and guiding principles

We suspected that the most difficult of these tasks would be describing the probabilistic time profiles of health states associated with the current and future years of life (Hamermesh (1985)). Based on early cognitive interviews it became clear that respondents thought in terms of experiencing specific illnesses. These were their unit of analysis of different risks in the current and future period(s). Respondents thought about *likely* illness "stories" or time profiles of health states they may experience over their lifetime. When respondents were asked to describe how they would experience their most likely illness, these stories had a latency period, a likely time of onset, a likely set of treatments, and a sense of the likelihood of recovery or premature death. The older the respondent, the more confident they appeared about both the likelihood of illnesses and their expected time-line of treatments and health states associated with each illness.

[3] Nearly all mortality risks involve a time period of morbidity, so the ideal study would measure the population's marginal rate of substitution between morbidity risks and mortality risks and income (i.e., all other goods). Studies that focus on only risks with low or no morbidity, or those that mistakenly omit morbidity risks, will obtain non-representative estimates of the population's marginal rate of substitution between income, morbidity risks and mortality risks.

Respondents believed some illnesses and illness profiles to be *more likely* than others based on their family history and current state of health. They undertook—or expected to undertake—interventions to reduce their risk of some illnesses but not others. In many cognitive interviews respondents recognized that to effectively mitigate most illnesses that would likely threaten them, they needed to adopt programs that reduced specific illness risks in their latter years of life rather than their earlier years.

As we began presenting prototypes of illness profiles to respondents, it became clear that respondents wanted to know with specificity the illness, its symptoms, the timing and duration of the illness as well as the prognosis. Respondents expressed consternation at illness profiles that they viewed as infeasible, vague or incomplete. Therefore, we spent the early portion of the design phase determining which of many possible attributes of future illnesses individuals cared about. Since we anticipated using a conjoint approach, we sought to identify the ten to twelve most important attributes that were common to the top ten to twelve causes of death or chronic diseases. We then searched for ways to consistently and clearly define these attributes and to communicate them to respondents (Baron and Ubel (2002); Moxey et al. (2003)).

1.1.2 Cognitive interviews and pre-testing

This process proved to be a great challenge that required "field-testing" many survey questions, graphics and formats. Over the course of nine months the survey went through four significant revisions. Due to the somewhat personal nature of many of the questions in the survey, we chose to evaluate prototypes of the instrument in one-on-one cognitive interviews rather than in a focus group setting. A principal investigator conducted each cognitive interview. These interviews began with the respondent taking the online survey as they would in the respondent's home, using a TV screen and keyboard. A PI remained present to answer and record questions of clarification and observe the respondent's behavior and attitude while taking the survey.

Once the respondent had completed the survey, a PI carefully debriefed the respondent by reviewing the survey modules and important questions, graphics or pre-designed response categories. Each interview lasted approximately one hour. We conducted 36 cognitive interviews over the study period.

We also pre-tested the penultimate version of the instrument on 142 respondents from Knowledge Networks' nationally representative panel. We then fielded it to a Canadian sample of 1,109 respondents in November 2002, which was drawn from an email list of approximately 4,000 addresses. While this sample would provide very useful information on Canadian demand for risk reductions, this also served as a second actual pretest for the survey instrument that we would use in the US. A third pretest was administered to a sample of 300 US citizens who were randomly selected members of the Knowledge Networks' nationally representative panel in early December 2002. The final version of the survey instrument was administered to a US sample of 3,000 respondents in December 2002.

1.1.3 Peer review of the survey instrument

During the development of the instrument, we drew on the expertise of a technical advisory board from fields including the psychology of risk communication, health economics, environmental valuation, and survey design. These technical experts evaluated the second of four

versions of the instrument. Each of these six experts provided extensive verbal and written feedback on the survey instrument. This panel included Victor Adamowicz, Richard Carson, Baruch Fischhoff, James Hammitt, Alan Krupnick and Kerry Smith. We are thankful for the invaluable experience, constructive criticism and advice that these experts have shared with us. Any remaining errors are, of course, our own.

1.1.4 Time and cognitive constraints

Meeting the time constraint of an expected thirty minutes for survey completion posed a considerable challenge. We, along with our technical advisory board members, had many more questions that would have been useful but would not fit into this timeframe. In addition, we needed to decide how much time to allow for the tutorial portion of the survey instrument. Our pretesting suggested that the quality of respondents' answers improved greatly with a careful tutorial that familiarized them with the metrics of each attribute of the illness profiles and the risk prevention program. Ultimately, we chose to devote over forty percent of survey time to the tutorial module (i.e. Module 2). The conjoint choice questions consumed about thirty-five percent of the survey time. The introductory and debriefing questions consumed the remaining twenty-five percent of survey time.

The quality of the preference information would be expected to decline if we exceeded the cognitive abilities of the respondents. This required us to consider very carefully the informational load that we imposed on respondents, the complexity of the three-way choices, as well as the cumulative fatigue and learning experienced by respondents. We sought to work within these constraints in several ways.

First, we methodically developed each respondent's familiarity with the choice process and the attributes of the objects of choice through our tutorial. Second, in light of individuals' limited ability to process complex risk information, we undertook two simplifications in how we represented illness profiles. To illustrate, consider an individual with a family history of early prostate cancer that tended to strike family members in their late forties. In reality, this individual faces a continuum of possible negative health outcomes: enlarged prostate only, treatable prostate cancer, prostate cancer that quickly metastasizes, etc. In other words, within prostate cancer, several distinct illness profiles are possible, each with its own subjective probability distribution of occurrence in each of future year of life. Our first simplification is that we ask individuals to make choices as if they faced only the one given illness profile for that illness.

Our second simplification is that we do not represent illness profiles as compound probabilistic events. When health researchers consider a concatenation of health states, they often first ask: "What is the probability of experiencing prostate cancer?" Then, conditional upon the type of occurrence at a particular age, health researchers describe the conditional probability of survival. We simplify the representation of a series of conditional events by describing a single marginal probability for that series.

Third, we explicitly evaluate the respondent's perceived level of complexity or difficulty for each choice exercise. After a choice opportunity, we ask respondents to rate the difficulty of that choice. While we discuss the details of these results below, the upshot appears to be that respondents became increasingly familiar with the choice process (and perhaps their own preferences) as the survey progressed. In addition, they did not experience measurable fatigue. Fourth, we have a rough proxy for the cognitive effort as measured by time devoted by each respondent to each portion of the survey. We retained the option to use these to screen out

choices made with such haste that the respondent could not have possibly read and processed the given information. For a more detailed discussion of fatigue and learning see Section 2.8.

1.1.5 The sample and sample selection

Correcting for as many sources of sample selection biases as possible is essential to ensure that our estimates of demand for risk reductions are truly representative of the U.S. population aged 25 and older. We expected to correct for three types of sample selection biases in the Knowledge Networks panel of respondents that we utilized. The first type of selection bias occurs when prospective panelists (over 525,000 random digit dialed households) are invited to join the panel but many decline the opportunity (at either the initial contact or in subsequent phases of the panel enrollment process). This is the most difficult and complex bias to correct for, requiring geographic information systems, telephone exchange spatial data, and a mix of US census data and other spatially indexed information.

The second type of sample selection bias occurs through attrition from the panel. After some period of participation, a panelist may drop out of the panel before we invite them to participate. Within the Knowledge Network setting, attrition can potentially be modeled using the wide range of profile data still available on each ex-panelist who has left the panel.

The third type of sample selection bias occurs when we invite one of the Knowledge Networks panelists to take a version of our survey and they decline. Often called non-response bias, this is typically the easiest bias to correct, because Knowledge Networks can readily supply sociodemographic information for all continuing panelists who were invited to participate in this particular survey. However, we are less concerned with whether the estimating sample represents the current KN panel than whether it represents the general population of the U.S.

We have thus undertaken an evaluation of the combined effects of all of these types of selection. A detailed description of this analysis is contained in Section **Error! Reference source not found.**. Somewhat to our surprise, we find that while selection into our estimating sample is systematic along several dimensions, individuals who are more or less likely to appear in our estimating sample differ mostly in terms of their marginal utility from avoided sick-time, and only slightly. In our models, because there are a number of influential outliers, we normalize this parameter on the overall *median* propensity for an RDD contacted individual to appear in our actual estimating sample.

1.2 Some specific modules of the survey

The survey is structured around four basic modules: (1) the risk beliefs and attitudes module, (2) the illness profile and risk tutorial, (3) the presentation of the choice sets, and (4) a debriefing and follow-up module. In the following subsections we refer to the form numbers that have been inserted in braces at the bottom center of each page in the single example of one of our survey instruments that is appended to the end of this document as Section 10. (The form number identifiers, of course, did not appear in the instances of the survey that were presented to respondents.)

1.2.1 Risk beliefs and attitudes

The survey opens with questions about risk beliefs and attitudes that encourage respondents to think about the environmental and illness-specific threats that they face (**Forms 2-8**).

1.2.1.1 Addressing the omission of substitute risk reductions

Many risk valuation studies do not identify the set of alternative risks that respondents face, nor do they measure the subjective level of these risks when valuing the targeted risk (Dow et al. (1999)). Therefore, they cannot describe how variations in individual-specific risk portfolios systematically affect demand for the targeted risk. In this section, we collect information that will identify the effects of this typically unobserved source of heterogeneity. We present respondents with the most complete set of health risks to date in a valuation study. This not only provides a more complete characterization of their choice set of possible risk reductions but also ensures that respondents are cognizant of potential substitute risk reductions when valuing the targeted risk reduction.

1.2.1.2 **Ex ante** *risk information and subjective risk levels*

Psychologists have shown that the salience of alternative sources of risk varies with individuals' information on these risks. These early questions also document respondents' experience with, and information on, each illness (**Form 3**). We also introduce and document respondents' knowledge of various states of morbidity (**Form 4**). Next we directly solicit a rating score describing how "at risk" respondents feel they are of experiencing each of these illnesses over the course of their lifetime (**Form 5**). In our empirical analysis, we can allow each respondent's answers to these questions to shift their marginal willingness-to-pay for a risk reduction. We interpret this "at risk" variable as a subjective attribute of the risks that respondents will consider in the subsequent choice exercises.

A notable feature of this section (and the entire survey) is that we present illnesses to respondents as distinct sources of risk. Many recent stated preference valuation studies have left the source of the risk vague, choosing instead to focus on a general and poorly defined risk of death (Krupnick et al. (2002); Chestnut et al. (2003)). In contrast, we have chosen to include all major illnesses and several important minor ones. These include: prostate cancer, breast cancer, colon cancer, skin cancer, lung cancer, heart disease (i.e., heart attack, angina), stroke (e.g., blood clot, aneurysm), respiratory diseases (i.e., asthma, bronchitis, emphysema), as well as diabetes and Alzheimer's. For subjective risk elicitation, we aggregated some illness labels based on the cognitive labels individuals used in our pretests. These included heart disease (i.e., heart attack, angina), stroke (e.g., blood clot, aneurysm), and respiratory diseases (i.e., asthma, bronchitis, emphysema). Aggregation was necessary to keep the list to a length that could be viewed comfortably on one computer screen. Each of these aggregated illnesses was described in greater detail in its illness profile later in the survey.

There are several reasons why we choose to include illness names. As we noted earlier, a major advantage of using these labels is that our pre-testing showed that individuals think in terms of specific illnesses when identifying hereditary risks and when planning for the mitigation of future risks. Second, the inclusion of the twelve major illnesses meant that our estimates of the marginal utility of avoiding a year of morbidity and premature mortality were broadly

representative of the leading lifetime illness risks. In addition, including diverse illnesses enabled us to motivate a wide range of health outcomes, (e.g., some associated with sudden death, such as heart attack and stroke, and others associated with chronic morbidity, such as diabetes and Alzheimer's disease). Gender-specific illnesses (e.g., breast and prostate cancer) are chosen to be consistent with the respondent's gender. Of course, the major disadvantage of specific illness names is that individuals may implicitly assume the presence of attributes that we did not explicitly include in the illness profile description. In empirical analysis, one could address this potential disadvantage by using illness-specific dummy variables to control for these effects.

Another difference between this survey and some other studies is that we chose not to give individuals extensive background information on each illness. Our primary reason for doing this is that we seek to estimate demand conditional on the individual's *ex ante* information set. We want to evaluate their *ex ante* preferences, not their updated preferences after being "educated" through the survey. Providing a primer on an illness is likely to give it more salience relative to those illnesses that the survey omitted. The option of providing a "primer" on each illness would have quickly overloaded the average respondent with information. It would also involve the opportunity cost of what else could be done with the limited amount of average panelist time available.

1.2.1.3 Addressing sequencing effects through randomization

Order effects may bias individual responses (Ubel et al. (2002); de Bruin and Keren (2003)). Therefore, we randomized the order in which we presented these environmental hazards and illnesses to respondents. For each individual, this randomly chosen order remained the same across **Forms 2-8** but it varied across individuals. In this way we sought to avoid order affects that might arise from either greater cognitive attention being allocated to the illness appearing first in the survey or from individuals inferring that the researchers viewed the first-ordered alternative as more important.

1.2.1.4 Potential for confounding by averting, defensive, mitigating behavior

The potentially confounding effects of averting, defensive, and mitigating behavior on demand estimates have been theoretically identified (Quiggin (2002); Shogren and Crocker (1999)). However, no empirical studies to date have attempted to identify and control for the effect of this behavior on demand. We endeavored to identify a subset of possible behaviors, their perceived relative costs, and their perceived effectiveness against specific illnesses.

The questions on **Form 6** explore the extent to which respondents feel they could *further* reduce threats to their health through a subset of changes in their behavior. This form is followed by the questions of how hard or personally costly (in terms of "time, money or effort") it would be to undertake these lifestyle or behavioral changes **(Form 7)**. The sequence of questions on these two forms helps us to distinguish between the respondent's understanding of their opportunity to control risks and their own personal subjective cost of doing so. Finally, individuals' propensity to undertake these behaviors will depend upon their perception of how effective these behaviors are in mitigating specific health risks. **Form 8** measures individuals' perceptions of how amenable each risk is to averting, defensive and mitigating behavior.

1.2.2 Illness profile tutorial

Sequencing the elements of the illness profile was a challenging aspect of survey design. We began this module by establishing the respondent's inter-temporal frame of reference. We reminded them of their current age and told them their expected age of death (**Form 9**) based on their personal characteristics.[4] We also informed the respondent that the rest of the survey would focus on health programs that would reduce their risk of getting sick and dying between now and their expected time of death. Throughout the survey, we conditioned the presentation of information on the respondent's age and gender.

1.2.2.1 Risk communication

Effectively communicating risk levels, and changes in those levels, to respondents is notoriously difficult in risk valuation studies (Corso et al. (2001); Fox and Irwin (1998)). We employed three approaches to communicate risk changes. First, in **Forms 10, 11, and 19**, we adapted and then augmented the risk-grid approach used by Krupnick et al. (2002). Visually, we represented a risk of 1 in 1,000 over the individual's remaining years of life expectancy. All colored squares represented the baseline risk, from which reductions would take place as a result of the intervention program. Although not visible from the attached black-and-white copy, the graphic represents the risk reduction by blue squares and the remaining risk in red squares.

To further illustrate and make the risk personal, we also represented the risk in its numerical form and presented its general nature textually in qualitative terms. For example in **Form 11** we present risk numerically as a mortality risk of 30 in 1,000 over forty years.

Third, we describe the percentage reduction of two risks from a common level in the choice sets. We included the percentage reduction for two reasons. First, it allowed us to address directly a common reasoning error described in the risk literature in which individuals only focus on the relative size of risk reduction (Fetherstonhaugh et al. (1997); Baron and Ubel (2002)). We took pains to point out cases where overall risk reduction was, in fact, very small even if the percentage (or relative) risk reduction looked large. A second reason for expressing the reduction as a percentage is that it may be used to directly compare two illnesses with the same baseline risk. The benefit of this approach is that it eliminates the need for the respondent to undertake two cognitive operations (e.g., subtraction and division) that would normally be required for careful comparison of the two programs. The only potential cost of this approach arises if the respondent rejects the conjecture that the hypothetical baseline risk for the two illness profiles would be the same. Not once did respondents raise this concern in cognitive interviews, while many said that the availability of the percentages facilitated their comparisons of the programs.

Finally, we directly warned respondents that they might overestimate the risk under consideration if they focused only on the numerical or percentage reductions. On **Form 19** we warn:

[4] In our one-on-one cognitive interviews and pretesting, we found that the typical respondent over-estimated his or her life expectancy by five to eight years, compared to standard age-dependent actuarial tables. Individuals frequently referred to their longest living relatives when answering our longevity question (**Form 44**). To prevent scenario rejection, we added eight years to our calculation of each respondent's life expectancy. This is a particularly important adaptation for respondents over the age of sixty.

> "Programs may be very effective at reducing your risk, but you should remember that your risks of dying may be very small.
>
> For example, consider a new program that reduces your risk of dying by 20%, from 30 in 1,000 to 24 in 1,000 over «XX» years. This may sound like a large percentage reduction, but your initial chance of dying was only 30 in 1,000 over the next «XX» years. To illustrate this below, the blue squares («blue square graphic») represent the size of this risk reduction. The red squares («red square graphic») represent your chance of dying even with the new program. («display relevant grid graphic»)

Risk reductions are thus represented in three metrics: a numeric representation, a graphical representation, and, finally, as a percentage of the baseline.

1.2.2.2 Risk comprehension

Following the risk portion of the tutorial module, we directly evaluate each respondent's ability to rank-order the magnitude of the two risks (**Form 20**).

1.2.2.3 Defining illness profiles to reduce omitted attribute bias

In **Form 12** we describe to respondents what we ultimately want to know from them, so that they understand why the information on the ensuing pages is relevant. We tell them we will describe how each illness might affect them, and then we will want to know which of the following two illnesses they most want to avoid (**Form 12**). They are told about two illnesses they might face and at what age these illnesses might strike. If they have already experienced one of the illnesses, they are asked to view the described onset as a recurrence.

Up to eleven attributes characterize each illness profile and program, although we concentrate on just the main attributes in most subsequent analyses. These illness profiles include the illness name, the age of onset, medical treatments, duration and level of pain and disability, and a description of the outcome of the illness. Our selection of these attributes was guided by a focus on those attributes that (1) most affected the utility of individuals and (2) spanned all the illnesses that individuals evaluated (Moxey et al. (2003)). In terms of the number and type of attributes, our design is comparable to existing state of the art health valuation studies (Viscusi et al., 1991; O'Conor and Blomquist (1997); Sloan et al. (1998); Johnson et al. (2000)).

In **Forms 14 and 15** we define the measure of morbidity that we use to describe these two illnesses. Adapting the types of pain and disability scales from several QALY indexes, we define for respondents what we mean by "moderate" and "severe" pain and disability. This provides respondents with a more concrete interpretation of these attributes as well as an understanding of the possible range of variation in each. We also introduce some types of treatments that are associated with morbidity. These include major surgery and minor surgery, as well as the duration of hospitalization, measured in weeks (**Form 15**).

We then describe for respondents the eventual outcome of each illness (**Form 16**). For the two illnesses under consideration in each choice set, we note there are four possible outcomes: (1) full recovery, (2) sudden death, (3) morbidity for less than six years with no recovery, followed by death, and finally by (4) chronic morbidity for more than six years, followed by death. For each of the two illnesses under consideration we describe the conclusion

of the profile in terms of the extent to which death is premature. We follow up this comparative information with a comprehension assessment to evaluate respondents' understanding of the information. This completes the introduction of the elements of the illness profiles.

Next, we introduce the interventions that could reduce the risk of experiencing these profiles. Most interventions took the form of medically driven risk management programs that centered upon an annual diagnostic test (**Form 17**). We chose this class of interventions because respondents viewed them as technically feasible and potentially effective. Respondents were familiar with comparable and pre-existing diagnostic tests such as mammograms, pap smears and prostate exams. Important from our perspective was the fact that this class of interventions could plausibly be applied to all of the illnesses upon which we focused. A second type of intervention (**Form 18**) involved the installation of new safety equipment in the respondent's vehicle to prevent the risk of injury in the event of an auto accident.

1.2.2.4 Minimizing payment vehicle bias

We sought to employ a payment vehicle that was: (1) applicable to most diagnostic programs, (2) generally accepted by respondents, (3) not confounded by too many other related benefits or costs, and (4) required multi-period payments. We sought this later property to emulate the continuing cost of actual public policies and private programs. Options for payment vehicles included changes in respondents' insurance premiums, higher government taxes in order to subsidize these tests, or copayments. Co-payments were the only vehicle that met the criteria described above. Copayments would have to be paid by the respondent for as long as the diagnostic testing and medication were needed. For the sake of concreteness, we asked the respondents to assume the payments would be needed for the remainder of his or her lifespan unless they actually experienced that illness. Costs were expressed in both monthly and annual terms.

1.2.2.5 Addressing concerns about hypothetical bias

If individuals' stated choices are affected by hypothetical bias, then their validity diminishes (Cummings and Taylor (1999); List (2001)). Hypothetical bias may arise from individuals having a strong incentive *not* to reveal truthfully their optimal choice. This bias may arise for several reasons. Individuals may strategically misstate their choice in hopes of manipulating the provision or future price of a public good. Alternatively, they may put little effort into seriously considering their budget constraint as they would in a real choice setting. Finally, they may wish to "please" the enumerator, leading to yea-saying. Scholars in the literature have explored three ways of mitigating aspects of hypothetical bias, all of which we incorporated into our survey design.

The first strategy is to include what is called a "cheap talk" reminder that encourages respondents to be cognizant of certain errors they might make because they are in a survey setting rather than a market setting. We sought to ensure that respondents recognize their tendencies to overstate their *WTP*, and to induce them to carefully consider their budget constraint (Cummings and Taylor (1999); List (2001)). The second strategy comes from the mechanism design literature which involves convincing respondents that their answers may actually effect the provision or pricing of the good under study (Carson et al. (2004)). The third strategy involves convincing respondents that there exist several acceptable and "good" reasons to reject the offer of the goods under study. This approach is intended to mitigate yea-saying or

respondents' inadvertent overstatement of their *WTP*. Discussing many of the legitimate reasons for opting out of the choice occasion also reinforces the role that economic reasoning should play in their decision making. It reminds respondents of the importance of substitute goods and binding budget constraints while compelling them to consider more carefully the relative expected value of the goods being offered.

We implement the three strategies to reduce hypothetical bias throughout our survey design. From the first screen we imply that respondents' answers may affect the provision of risk mitigating programs (Carson et al. (2004)). **Form 1** states: "Your answers may help public officials provide you with better ways of managing your health." We further develop this context on **Forms 17, 20, 21, 22, 24, and 27**.

Second, in an effort to mitigate hypothetical bias we include versions of a cheap talk script (Cummings and Taylor (1999); List (2001)). **Form 22** begins "In surveys like this one, people sometimes do not fully consider their future expenses. Please think about what you would have to give up in order to purchase one of these programs. If you choose a program with too high a price, you may not be able to afford the program when it is offered…"

We then focused respondents' attention on their option to choose neither program. In an effort to mitigate yea-saying, by dispelling the respondent's assumption that they were "supposed" to choose one or the other program, we listed four plausible and legitimate reasons for why a reasonable person might reject both programs, choosing instead the "neither" option (**Form 22**). "We give you the option of choosing neither program. People might choose neither program because they:

- could not afford either program,
- did not believe they face these illnesses or injuries,
- would rather spend the money on other things, or
- believe they will be affected by another illness or injury first."

As a final check on a particular subset of the reasons for hypothetical bias, we directly asked respondents if they felt they could actually pay for the programs they had chosen. Of course such a question would not test for, or reveal strategic behavior, since presumably they could answer this question strategically as well. However, it elicits the respondent's assessment of their own intended purchase behavior, thereby revealing whether he or she feels they have made carefully considered and realistic choices.

1.2.2.6 The duration and effectiveness of the risk programs

Before presenting respondents with choice sets, we sought to ensure that respondents clearly understood the intertemporal range of program benefits. The survey described (for two illnesses) the time of onset, time of death relative to their expected lifespan, the baseline risk and risk reduction (**Form 23**). On the same page, the survey said, "We want to be clear about when the benefits of each program begin. For example, the benefits of Program A are that your risk of illness A is reduced from X in 1,000 to Y in 1,000, starting when you are ZZ years old and continuing for the rest of your life."

We also focused the respondent's attention on the *status quo* option if they chose neither program. Recall that we have already elicited respondents' beliefs about what illnesses they are most likely to experience over their lifetimes in the absence of these risk reducing programs (**Form 5**). The survey stated, "If you DO NOT choose Program A, your risk of illness A will remain at X in 1,000 over this time period" (**Form 23**). Prior to the choice questions, the survey

stated, "If you choose neither program, remember that you could die early from a number of causes (of death), including the one described below" (**Form 25**).

We endeavor to counter another “survey effect” that may arise if individuals are skeptical of the stated effectiveness of the programs. We did this by directly acknowledging the survey context in which the respondent was to make their choice (just as the cheap talk language does). Furthermore, we acknowledge that it might be reasonable for individuals to be uncertain or skeptical of the stated effectiveness of these programs. Having identified and acknowledged this potential bias we then ask them to make their choices *as if* they had been shown proof that the programs performed as described in **Form 24**.

2 Stated Preference Quality Assurance and Quality Control Checks

Equation Section (Next)

In this section, we go into greater detail on several points that are relevant to ensuring that stated preference data are of sufficient quality to warrant the use of results based upon them for policy analysis.

We undertook numerous *ex ante* measures to minimize biases through careful survey design and also seek to evaluate our data *ex post* for the presence of remaining biases. Our survey includes a verification of respondents' risk comprehension well as features to limit the extent of biases associated with the hypothetical nature of the choice questions, distortions due to the omission of relevant substitutes, order effects across the choice questions, and yea-saying tendencies. Our choice set design is structured to provide ample opportunity for external "scope" testing, as well as for general evaluation of the validity of results in relation to economic theory.

We also include in this section some information in response to the concerns of reviewers of previous versions of the main paper. In particular, we discuss respondents' potential use of choice heuristics and their potential recoding of attributes, and whether respondents to our survey can be assessed with respect to the consistency of their choices.

2.1 Risk comprehension verification

After we administer an extensive risk tutorial and present the risk changes in three forms (textually, graphically and mathematically), we test the individual's risk comprehension. This comprehension test requires individuals to rank the sizes of the risk reductions associated with two risk mitigation programs. Approximately eighty percent of the individuals demonstrated adequate comprehension of the relative risk size reductions of the programs, which is a rate consistent with risk comprehension levels documented in other surveys (Alberini et al. (2004) and Krupnick et al. (2002)).[5]

2.2 Minimization of biases associated with omitted substitutes

In contrast with many valuation studies that focus on just one or two risks and their associated risk-reduction programs, we endeavor to reduce biases associated with so-called bracketing (Read et al. (1999)) via inclusion of nearly all major competing health risks (and specific programs to reduce them) across each individuals' choice sets.[6]

Presentation of a broad spectrum of major health threats and mortality risks increases the generality of our estimates. Of course, a potential disadvantage of this approach is the cognitive

[5] As Harrison and Rutstrom (2006) argue, reliable estimates of the monetary value of risk reductions hinge on respondents' comprehension of mortality risks. Their research suggests that it is indeed possible to elicit subjective beliefs about mortality risks from individuals. We do conduct a sensitivity analysis of the effects (on the estimated parameters) of including and excluding individuals from the sample based on their risk comprehension. A priori, we cannot expect people to make rational choices if they do not understand the simple concept of risk upon which our survey's choice questions rest, so we do not include these individuals in our estimating sample. However, our auxiliary sensitivity analysis demonstrates that inclusion of respondents who fail the risk comprehension test does have an effect on our parameter estimates, so this exclusion decision is important.

[6] Ashenfelter and Greenstone (2004) also address the problem of omitted variables and other biases in measuring the value of a statistical life. Competing risks are addressed in Dow et al. (1999).

complexity associated with the choice task, which we seek to minimize through careful survey design, and which we evaluate *ex post*.[7]

2.3 Minimization of hypothetical bias

At the beginning of the valuation module, to minimize hypothetical bias, we include a "cheap talk" reminder—to ensure that respondents carefully consider their budget constraints and to discourage them from overstating their willingness to pay (Cummings and Taylor (1999); List (2001)). Individuals are instructed, "In surveys like this one, people sometimes do not fully consider their future expenses. Please think about what you would have to give up, to purchase one of these programs. If you choose a program with too high a price, you may not be able to afford the program when it is offered…."[8] The second strategy comes from the mechanism design literature which involves convincing respondents that their answers may actually effect the provision or pricing of the good under study (Carson et al. (2004)).

2.4 Minimization of distortions from provision rules and order effects

To clarify provision rules for each choice set (Taylor et al. (2010)) and to avoid potential choice set order effects (Ubel et al. (2002); de Bruin and Keren (2003)), we instructed individuals to assume that every choice is binding and to evaluate each choice set independently of the other choice sets. Our empirical analyses show that the first four choice sets appeared largely free of choice task order effects. Individuals did exhibit a slightly higher propensity to select a program from the last choice set, an effect that has also been demonstrated in other similar settings (Bateman et al. (2004)).

2.5 Tests for the effects of program scope

We explore whether individual choices are sensitive to the scope of the illness profile and the scope of the risk mitigating program (Hammitt and Graham (1999); Yeung et al. (2003)). We show, even in the simplest possible choice models, that individuals readily pass the "scope test." Our subjects are highly sensitive to differences in the scope of our key choice-scenario attributes across the 7520 different choice scenarios considered by our 1801 individuals. Even an extremely parsimonious conditional logit choice model, specified in terms of a minimal number of raw program attributes, produce intuitively plausible and strongly significant coefficients on the two most crucial aspects of each program: i.e. a lower cost and a greater risk reduction make a program more attractive. When we add the other two most important dimensions of the illness profiles—the number of sick-years and the number of lost life-years for which the risk will be reduced—these are shown also to be strongly significant determinants of respondents' choices among programs. Respondents are systematically more likely to choose programs which address more serious health threats.

[7] We assess this concern directly in the survey. After each choice set we ask individuals how difficult each choice was. On a scale of 1 to 5 (very easy to very difficult), the average response for the first choice set was 3.2. This rating fell with each subsequent choice set, suggesting that the choice task became easier with increasing familiarity.
[8] For a complete description, see the annotated survey instrument available from the authors. We note that Hakes and Viscusi (2007) have demonstrated that the value of a statistical life implied by stated preference survey estimates is not statistically significantly different from estimates of the same quantity derived from seatbelt usage.

2.6 Minimization of yea-saying

Another concern, if there are no actual costs to respondents at the time they agree to purchase a hypothetical good, is that they will "yea-say," that is, agree to purchase the offered good in a effort to be agreeable. We employed a strategy that involves reminding respondents that there exist several "good" and acceptable reasons to reject the offer of the goods under study.[9] This approach is intended to mitigate yea-saying which may lead to the respondents' inadvertent overstatement of their WTP. Discussing many of the legitimate reasons for opting out of the choice occasion also reinforces the role that economic reasoning should play in their decision making. It also reminds respondents of the importance of substitute goods and binding budget constraints while compelling them to consider more carefully the relative expected value of the goods being offered.

2.7 Basic tests for theoretical validity

An important test of the validity of individuals' stated choices is whether their WTP varies with specific variables as theory would predict it should. In a variety of statistical analyses that make use of these survey data, we have shown that respondents' stated WTP does vary systematically with their income, age, discount rate, and health status. It also varies (in directions that theory or intuition would predict) with the latency, duration and severity of the illness profiles as well as the cost and the effectiveness of the program as measured by the size of the risk reduction.

2.8 Respondent learning and fatigue

In response to the number and complexity of choice tasks, respondents may both learn and become fatigued. Learning about both their own preferences and how to more efficiently choose might reduce the amount of time respondents spend on each choice task. Increasing fatigue, in contrast, may increase their time-on-task. These processes are important for us because learning might reasonably be expected to increase the quality of preference information we can recover from their stated choices, while fatigue might reasonably reduce it.

We evaluate these effects in three ways. First, after each choice set we ask individuals about the subjective difficulty of that choice, using a rating scale for difficulty. (See the single example of one instance of our survey appended to this document as Section 10.) On a scale of 1 to 7 (from "easy" to "very difficult"), the average response for the first choice set was 3.2. (See **Forms 26**, **30**, **34**, **38**, and **42**.) We asked respondents to continue to rate the difficulty of each of their choice tasks. The first such subjective rating can be expected to be fairly arbitrary, since the respondent must decide for themselves "relative to what?" However, these difficulty ratings, on average, tend to fall with each subsequent choice set, suggesting that respondents perceived that the choice task became easier with increasing familiarity or learning.

[9] Just prior to the introduction of the first choice set, recall that the survey volunteered that "People might choose neither program because they:

- could not afford either program,
- did not believe they face these illnesses or injuries,
- would rather spend the money on other things, or
- believe they will be affected by another illness or injury first."

Second, we examined trends in the amount of time respondents spent on selecting an alternative from each choice set. On average, respondents' time spent on each choice task fell consistently from their first, second and third choice tasks, thereafter remaining relatively constant across their fourth and fifth choice tasks. This result, which is consistent with respondents' self-reported difficulty ratings, suggests that fatigue at least did not slow them down.

Third, we explored whether there was systematic variability in individuals' *WTP* across the five choice tasks. One might expect that if individuals were becoming fatigued, their answers would become increasingly random. Increasing randomness might also occur if individuals were "rushing" through the choice tasks. (This concern might be heightened in light of the above evidence of progressively falling time spent on each subsequent choice task.) Our test for trends in implied *WTP* as function of the order of the choice tasks show no discernable trend up or down. The only pattern that clearly emerged was a slight increase in *WTP* for the very last (i.e., fifth) choice task. Respondents were informed between the fourth and fifth choice tasks that the next choice would be the last program choice they would be asked to consider, so this could be a "home stretch" phenomenon.

2.9 Heuristics and metric recoding

Respondents' use of heuristics in decision making is indeed a very important consideration and one to which we devoted a great deal of care to minimize and evaluate through the many iterations of trial versions of the survey format with 36 survey test subjects over several rounds at the Knowledge Networks facility in Menlo Park, CA. But it is important begin by asking what is a fair standard for the "eligibility" of preference data and whether expectations for stated-preference (SP) data, as opposed to revealed-preference (RP) data, may represent a double standard. Would RP data be disqualified from use if it were found to be affected by heuristics? The entire field of behavioral finance and a growing number of influential field experiments suggest that answer is clearly no.

Given this, the next important question is whether SP data is more likely to be affected by heuristics that comparable RP data would be. First, our respondents probably see more information, more comparably presented, than they would be shown to them in any real choice situation with respect to opportunities to reduce risks to their lives and health. Moreover, we probably spent more time, and provide more learning strategies (with risk measures and graphics) to prepare them for their decision making than would many medical office visits where patients at their annual check-ups must consider their doctor's recommendations to elect (and subsequently pay for) a variety of diagnostic tests.

The next concern is whether respondents selectively discard or recode the information they are given and report their preferred alternative from each set in a way that renders these choice data unusable for the purposes of recovering an informative estimate of their *WTP* for health risk reductions. As experimentalists, we must first to acknowledge that it is not possible to observe directly any individual's mental decision process, so we likewise cannot observe the presence or absence of heuristic processes. What we can do is to look for evidence that such heuristics, to the extent that they exist, have damaging consequences for our data and the inferences we draw from them.

Do we see any blatantly obvious evidence of the use of damaging heuristics? No. First, the attributes of the illnesses, and the characteristics of the respondents, are strongly statistically significant predictors of choices. Second, the nature of the systematic variability that we identify

is consistent with what intuition and general economic theory would predict. Third, the fitted *WTP* amounts based on our final SP estimates of consumer preferences are generally consistent with the available benchmarks for RP data that exist within the literature.

2.10 Concerns about choice inconsistency

One concern is that due to the complex nature of the choices individuals may not correctly and consistently evaluated the risk-tradeoff questions. If this is true, we would observe respondent failing internal consistency checks such as those for transitivity of preferences. If the complexity overwhelmed respondents such that their choices did not preserve the properties of transitivity then a degree of randomness would characterize the choice data. In the extreme, if this happened, the observed choices would appear predominantly random. Our estimates of the marginal utility parameters for different illness profile attributes would not be statistically significant. We conclude that choice inconsistency does not appear to be happening in the extreme and certainly not so much as to prevent us from getting fairly precise measurements of the central tendency of preferences.

2.10.1 Reed Johnson's VALIDTST program

In some simpler choice contexts, where there are no more than ten different levels for each attribute, and where utility can be assumed to be either strictly increasing or strictly decreasing in each distinct attribute, it is possible to assess choice consistency systematically, using a Gauss program called VALIDTST.PRG, prepared by Reed Johnson. This program, written in 2004 (subsequent to the fielding of our survey), is designed to take conjoint choice data and test for "stability, monotonicity, transitivity, and dominance relations in SP designs."

The VALIDTST program allows the researcher to specify the number of attributes and the number of alternatives as well as the number of choice task repetitions. The researcher must specify whether each attribute has levels with are decreasing (-1), increasing (+1) or unordered (0). Identifiers must be provided for each respondent and for each choice. Unfortunately, the program appears to allow for no more than ten possible levels for each attribute, since each attribute level must be represented by a single digit, to be concatenated into a string. Many of the attributes in our choice sets have far more than ten possible levels.

The different available tests in the VALIDTST program are as follows (with discussion concerning the appropriateness to our study appended in each case):

1. Look for stability relations in repetitions of the same choice: A~B~C, A~B~C: "If A1=A2, B1=B2, and C1=C2 then Choice1 must equal Choice2"
    - Our survey design involved random draws of attribute levels for each alternative in each choice set (as described in Section 3). Given that choices can be identical only for individuals of the identical age and the same gender, and given the number of attributes and the number of possible levels for each attribute, it is highly unlikely that there is ANY pair of identical choice sets anywhere in our 7520 choice sets, let alone among the five choice sets posed to any one individual.

2. Look for within-pair monotonicity: A~B/A~C, where all elements of B <= C: "If A1<=B1, then Choice1 should not be A"

- Our survey design rejected any pair of program alternatives for which one program both cost less and produced a greater risk reduction than the other. Since we are confident that preferences are probably monotonic in cost and in the sizes of the risk reductions, there are no cases in our design where we could test for within-pair monotonicity.

3. Look for across-set monotonicity: A~B1~C1, A~B2~C2: "If B1<=B2, then Choice1 = B1 does not allow Choice2 = A"
    - Each choice set presented to an individual involves a different disease label. For each gender, there were eleven possible disease labels, and we used ten labels randomly selected from this set, sorted randomly into pairs, for each of the five choice sets. In no case would the same individual see the same alternative paired with two different alternatives in two different choice sets.

4. Look for consistency relations: A~B~C, A~B~D: "Choice sets A B1 C1 and A B2 D2. If B1=B2, then Choice1 = B1 does not allow Choice2 = A, and Choice1 = A does not allow Choice2 = B2"
    - Again, there is no instance in our survey design where the identical alternative A is paired with two different alternatives B1 and B2 for the *same* individual. Across all of the potentially 15040 unique hypothetical programs used in the 7520 different choice sets faced by respondents in our estimating sample, each alternative is a random combination of ten or eleven different attributes, most with a very large number of possible levels. It is very likely that there is no pair of identical alternatives A proposed to any pair of individuals, either.

5. Look for transitivity relations: X~Y, Y~Z, X~Z: "For unique profiles X, Y, and Z, if Choice1=Y, Choice2=Z, then Choice3 can't equal X"
    - In our survey, the same person never sees the same alternative X in more than one choice set. All ten of the alternatives seen by each respondent are different. In particular, a different disease label is associated with each of the ten non-status-quo programs described in the five choice sets presented to each individual. Thus it seems we have no real opportunity to test for transitivity for any given individual. Heterogeneity in subjective risks of different illnesses or injuries will mean that we cannot expect identical preferences across individuals, so even if we could find some set of identical alternatives posed to two individuals, there can be no expectations that they would make choices that imply transitivity. We estimate average preferences in contexts where we do not control for systematic heterogeneity in preferences.

6. Look for dominance (no-tradeoff situations): A~B/A~C, A~B~C, where preferred attribute is z: "Choice sets A1 B1 C1 and A2 B2 C2. If A2(x,y) > A1(x,y), B2(x,y) > B1(x,y), and C2(z) < C1(z), then Y = 1."
    - As mentioned above, we rejected all choice sets where the two substantive alternatives exhibited strict dominance in terms of program cost and the size of the risk reduction. Thus, by design, no substantive pair of alternatives involves strict dominance and both of the substantive pairs involve a greater (e.g. non-zero)

risk reduction but a higher cost (lower net income) than the status quo. Thus there seems to be no opportunity to find any such case in our data.

Thus, unfortunately, we seem to have too many attribute levels to make use of this helpful utility, and likely no more than a very tiny fraction of cases, if any, where it would even be possible to assess these problematic relationships. Also, we know that preferences with respect to lost life-years, for example, depend empirically upon the number of prior sick-years, so that lost life-years, if an illness is bad enough and long enough, could in principle be viewed as a good thing (from the perspective of the present). Thus it is not even possible to assert, unambiguously, that increasing levels of some attributes imply increasing utility. Preferences are not strictly monotonic in every attribute (and in any case, utility, empirically, is not linear and additively separable in each attribute).

With much larger samples of people in each of our 135 unique age/gender "bins", we might have considered how to seed our choice sets with cases that would catch occasional problems in terms of choice consistency. However, ex ante, we were concerned that we should have enough independent variation in attribute levels across choice sets and across individuals so that we could identify robustly statistically significant marginal effects of illness profile attributes on indirect utility levels. Of course, there is always room for follow-up research with our survey instruments, and new samples of people, wherein one might conduct specific tests of the extent to which some individuals show evidence of the types of anomalies assessed by the VALIDTST program.

2.11 Unobserved recoding and scope insensitivity

A second concern is that even if choices are not random, respondents may recode the risk information from the metric in which it is given to them into one that is simpler and unknown to the researcher. It was this type of recoding and the resulting insensitivity to the scope of good that led to the recommendation for mandatory "external scope tests" in early CV formats. The proper way to conduct this external scope test is to "split the sample" and to administer one amount of the good some respondents and a different amount to others. The objective is to demonstrate that different people are willing to pay different amounts for different quantities of a non-market good. If respondents are prompted, in the context of the *same* questionnaire, to evaluate both a larger and a smaller amount, the test is merely an "internal scope test."

We undertook a variant of the required scope test that is analogous to splitting the sample in the way we designed our choice sets. Importantly, the choice scenarios vary across respondents, as well as across choice sets. For each of the 1801 individuals whose choices were used to produce our estimates—indeed, for all 7520 choices analyzed in our data, the illness profiles were essentially unique.

Had *all* respondents seen the same five choice sets, we would not have met the external scope test requirement. We would only have been able to show that WTP was larger when benefits were larger "between scenarios." This would have been a weaker internal scope test. However, instead of splitting our sample of 1801 individuals into two groups, and showing everyone within the same group the same set of choice scenarios, we effectively had 1801 different groups. We contend that this strategy actually vastly outdoes the usual "external scope test" because every respondent considered different illness profiles and different risk-reduction

program costs. We did not merely split the sample into *two* groups, each of which saw a different level of benefits and different costs.

3 Details of the Choice Set Design

Equation Section (Next)

A key feature of the "experimental design" of the choice sets used in this study is that every choice set posed to every individual is essentially unique. This is different from the usual case in conjoint choice experiments. Many conjoint choice scenario designs involve a limited set of survey instruments (different "versions") where the mix of attribute levels across versions is designed to insure sufficient independent variation in attribute levels to permit estimation of "main effects" (and sometimes higher-order effects). When every respondent to a survey is eligible to receive each different design, blocked designs can be used to improve estimation efficiency. In this study, however, there are strongly binding constraints on our ability to consider blocked designs. The illness profiles that can be offered to a respondent depend crucially on the individual's age at the time of the survey. The types of illnesses which can be included in the individual's questionnaire depend upon the individual's gender.

3.1.1 Rationale for our approach to randomization

In the typical conjoint choice experiment, where there is a relatively small set of survey versions, it is often *possible* for the researcher to assign one randomly selected version of the survey to each respondent in a particular study. This strategy insures that each survey version is given to some large and approximately equal number of respondents. Here, however, the range of eligible choice scenarios for each individual must be indexed to the respondent's specific age and gender. Illness profiles relevant to a 25-year-old are not relevant to a 75-year-old. Likewise, some illness profiles relevant to women are not relevant to men (and vice versa). We have 135 different age/gender combinations in our sample. While it may have been possible to treat each age/gender combination as a separate subsample, and to tailor a fixed set of illness profiles and testing programs to be assigned randomly within that group, the time costs of doing so seemed prohibitive (and many of these groups were simply too small—see Table 3-1 for the complete joint distribution of age and gender in our sample). This is why we opted for individual randomizations rather than any attempt at a standard blocked design.

If we had been willing to give up on using a representative sample from the population, it may have been possible to sample intensively from just a subset of ages for each gender and thereby to acquire sufficient people in each of several age/gender bins to permit a formal blocked design. Within each bin, however, this would also have limited the number of "treatments" in the design, and we placed a premium on being able to represent the widest possible range of illness profiles and illness labels, where an illness profile involves an arbitrary partitioning of the individual's remaining life expectancy into a set of up to four mutually exclusive and exhaustive intervals in different health states.

In stylized marketing experiments, it is often feasible to assign choice sets without regard to the attributes of the research subject (e.g. age and gender), but in this study, the choice sets are fundamentally dependent upon the individual's age and gender, so the options for the experimental design behind the assignment of choice set to individuals is greatly restricted. The large number of necessary attributes, including our desire to assess

a dozen different illnesses within a common framework, and the fact that utility is highly likely to be non-linear in these attributes, means that many other familiar simplifications from the experimental marketing literature are unavailable in this context. For example, there are few opportunities to use common designs or common choice sets across people. While an experimental marketing framework permits the type of spare and elegant design that makes it easy to discern greater subtleties in that stylized context, some tradeoffs must necessarily be made in an ambitious attempt to understand a much more complex and realistic choice context with a variety of health threats, long and variable time horizons, and a huge amount of natural interpersonal heterogeneity that may be potentially relevant to choices.

3.1.2 Framework to permit an "external" scope test

When each respondent considers multiple choice sets, and when few enough attributes with few enough levels are being considered, it may be possible in some cases to use the identical survey instrument for all respondents in a given group. However, the standard "scope test" in stated preference applications requires that *different* respondents be asked to consider choice scenarios with different levels of "benefits." This requires at least two versions of the survey instrument. In this study, however, we have as many versions of the survey instrument as there are respondents. Each survey instrument is constructed based on the same template, but all of the relevant, uniquely randomized quantitative information concerning each illness profile and each corresponding risk reduction program is converted to character strings and merged into the survey template on a respondent-by-respondent basis. Thus, we have pushed the standard scope test to its logical extreme.

3.1.3 Outline of the choice scenarios

Each respondent receives five choice sets consisting of two different illness- or injury-reduction programs plus a "neither program" option. Ten different illnesses are therefore represented, out of a total of eleven possibilities for their particular gender. No risks are repeated across the five choice sets faced by each individual, and the composition of each pair is random. The only restriction on the order of the names assigned to each illness is that for the first choice set, the attributes of which are also used in the tutorial portion of the survey, the choice scenario cannot include the "traffic accident" risk, since the story surrounding this type of risk is different than for the disease risks.

For each respondent, age and gender are known in advance, so five randomized conjoint choice sets can be generated specifically for someone of that age and gender. No respondent is asked to consider an illness or injury that is described as first affecting them at an age younger than they are at present; males are not asked about breast cancer (even though a tiny fraction of breast cancers cases are male), and females are not asked about prostate cancer. However, the sum of breast cancer and prostate cancer alternatives is roughly equal to the number of instances of each other single type of illness or injury employed in the choice scenarios. The frequencies for each type of illness name are given in the first row of Table 3-2.

To the extent that individuals feel that the assertions about a particular illness in their choice scenarios do not match their subjective illness profiles in those cases, our

follow-up questions for each choice task, about the extent of this correspondence, can be used to correct for any mismatch.

3.2 Nominal life expectancies

Life expectancies are normally based on current estimates, by year of age, independent of gender. For each respondent, current age can be used to assign the appropriate actuarial life expectancy. These life expectancies include all causes of death, whether the individual dies "prematurely" or otherwise.

For this study, however, we desire a hypothetical life expectancy that is *plausible* for each individual respondent. In pretesting, it became clear that subjects, when considering their life expectancies, tended to assume a life expectancy conditional on the assumption that they would not *"die early."* This strategy stems from an apparent tendency of focus-group members to pay attention only to the age at death of grandparents and/or parents who did *not* "die early."

We thus add an arbitrary eight years to each actuarial life expectancy, since this seemed to be sufficient to preclude rejection of the life expectancy assertion for most pretest subjects. We then permute the stated life expectancy for women by adding one year, and that for men by subtracting one year. These adjustments are arbitrary, but they helped to overcome respondents' frequent tendency to reject their official actuarial life expectancy. We frame each illness in our stated-preference choice scenarios in terms of its hypothetical independent effect on this "no-premature-death" life expectancy. In estimation, we have the opportunity to correct for any mismatch between stated and subjective life expectancies by using respondents' answers to our follow-up question about their individual subjective life expectancy.

3.3 Specific stylized illnesses and injuries

A degree of contrivance is necessary to achieve sufficient independent variation in all of the attributes of different illnesses or injuries that may threaten the individual's life or health. In principle, it is the attributes of a disease alone that should determine the individual's willingness to pay to avoid it, since these prospective experiences should have a direct effect on expected utility levels.

However, we suspect that the *label* attached to a disease may have a systematic effect on *WTP*, beyond the short list of objective illness profile attributes we describe to the respondent. For example, the U.S. EPA has been concerned for some time that there may be a "cancer premium" that needs to be associated with *WTP* for reductions in cancer-causing health risks. Thus, we employ eleven different possible labels to the diseases or injuries for which our conjoint choice scenarios offer risk reductions:

- a gender-specific cancer (breast cancer for women, prostate cancer for men)
- colon cancer
- lung cancer
- serious skin cancer
- heart attack
- heart disease
- stroke

- respiratory disease
- diabetes
- Alzheimer's disease
- traffic accident

Our survey begins with the unfolding of the different attributes across which respondents will be comparing each pair of diseases, and the programs that are offered to reduce these risks. Again, for the initial choice pair used in this phase of the survey, we exclude traffic accidents, since the story is a little different for traffic accidents than for the typical disease.

3.4 Reductions in lifespan due to non-fatal cases

For illnesses and injuries which are not fatal (from which the individual recovers within their stylized life expectancy), we nevertheless describe one of the consequences of having had this illness as being a potential decrease in life expectancy, perhaps from subsequent greater vulnerability to other health threats. For example, the individual may recover, but "die at 89 instead of 92." These reductions in life expectancy are randomized by calculating the time interval between the age at recovery and the nominal life expectancy, and then shrinking it by a factor randomly drawn from the following list: 0.00; 0.05; 0.10; 0.15; 0.20; 0.25; 0.30; 0.35; 0.40; 0.50.

The algorithm is set up to allow different eligible lists for each disease. One potential problem is that younger people will, on average, be told of larger average life expectancy reductions due to the same disease. It is possible that having the same disease later in life may compromise your life expectancy more, but this is uncertain. Despite the correlation between current age and the size of the reduction in life expectancy, there remains considerable independent variation. Still, our basic models diligently control for any effects of respondent age at the time of the survey.

3.5 Risk descriptions

Risk reductions due to each program are described in terms of the baseline risk and the new risk, as well as the percentage risk reduction that this difference represents. To avoid confounding the apparent risk reduction in percentage terms by using different baseline risks, this baseline risk is constrained to be identical *within* each choice pair. These lifetime baseline risks are drawn randomly from a universal list of "risks per 1,000" over their remaining life. This list includes: 4; 5; 8; 10; 15; 20; 30; 40.

Risks are cast in terms of the chance in 1,000 so that a conventional 25-by-40-cell grid can be used to depict the (small) absolute levels of risk. It would be preferable to be able to convey the actuarial risks of each illness as a function of age in a two-dimensional graph, and to depict risk reductions as a shift in this age profile. Likewise, it would have been preferable to depict separately the risk of incidence and the risk of death, conditional on incidence. However, this would have required two graphical representations for each illness. A competing need is to have each choice scenario be completely described in the minimum amount of space—ideally in just one computer screen in legibly sized fonts. Focus groups also determined that the ability of the average respondent to interpret graphical information was unfortunately limited at best (at least in an environment without an interviewer to help with this interpretation). We thus opted to

describe the compound risk of incidence and mortality, giving up some realism in the description of the risk in exchange for compactness. We employ "representative" illness descriptions (latency, symptoms, recovery/death) and ask respondents to view these as the expected trajectory of the illness. Thus, we planned to be able to disentangle incidence from mortality since there are both fatal and non-fatal versions of most of the illnesses and injuries we cover.

For each baseline risk, there are a number of possible reduced risks. These reduced risks are limited to levels that correspond to "round-number" percentage reductions. See Table 3-3. The randomization process is set up to guarantee that no pair of programs will be characterized by the identical risk *reduction*, even though each alternative in any given pair has the same *baseline* risk.

3.6 Conceptualization of risk changes and illness profiles

In developing our illness profiles, it was clear right away that the full complexity of the range of future health states faced by an individual would have to be simplified. Even if we limited our attention to illness profiles that included just a pre-illness period of current health, some sick-years (or fractions thereof), some potential recovered or remission years, followed by some potential lost life-years, there are four spells to take into account.

With four spells, given the individual's current age and nominal life expectancy, three of these spells could take on any length greater than zero as long as no period exceeded the individual remaining lifespan and as long as the sum of these periods was less than or equal to this remaining lifespan. The other period would be defined as the remaining lifespan minus the sum of the other three spells.

Even if we could assume that time in each health state was homogeneous, this would mean that each illness profile would be a point on simplex in four dimensions with each vertex a distance from the origin defined by the individual's remaining lifespan. For any given type of illness, there would be a density function defined on this simplex. For example, most cases of a fatal illness that strikes late in life will have long latencies, short spells of sickness, no recovery, and few lost life-years. There might be a joint expectation of this distribution of illness profiles, but considerable variability in the form of variances and covariances among the four different durations, constrained by the adding up requirement.

It is too difficult to visualize a four-dimensional density, so we illustrate the development of our survey's characterization of illness profiles by simplifying the story to a two-dimensional simplex. Imagine an individual with 50 years of remaining life, and consider an illness risk that involves only sick-years and lost life-years (i.e. we will assume latency and recovered/remission time is always zero).

Suffering this illness is not certain. Without intervention, the individual has a baseline probability of getting sick that equals Π^{NS}, and a probability of remaining healthy (i.e. of experiencing zero sick-years and zero lost life-years) of $1-\Pi^{NS}$. Figure 3-1 illustrates a specified illness risk, along with the different illness profiles that could go along with this illness (in the case with just sick-years and lost life-years). We sketch a

bell-shaped distribution, centered around twenty sick-years and thirty lost life-years (i.e. this is a serious illness).

In our survey, we limit the types of programs offered to those which merely change the probability of suffering the illness in question, as in Figure 3-2. If the individual purchases the risk-reduction program, their chance of remaining healthy (sick-years = lost life-years = 0) is increased to $1-\Pi^{AS}$ and their chance of getting sick is reduced to Π^{AS}. However the shape of the distribution of illness profiles, conditional on getting sick, is unchanged. All that happens is that the density function associated with the mix of sick-years and lost life-years is scaled down, with the reduced probability being added to the probability of staying healthy.

Notice that our characterization of health risk programs precludes another possibility. As in Figure 3-3, it could be the case that the program does not affect the probability of getting sick, so that $\Pi^{AS} = \Pi^{NS}$. What happens instead is that the course of the illness is changed. In Figure 2-3, the program causes the distribution of sick-years and lost-life years to shift, so that the distribution of illness profiles is characterized by more sick-years and fewer lost life-years. In this limited scenario, the program doesn't prevent people from getting sick, it just keeps them alive longer. Program effects such as those illustrated in the two-spell case in Figure 3-3 are not considered in our choice scenarios.

Our cognitive interviews with test subjects made it clear, early on, that relatively few potential respondents were comfortable with diagrams like Figures 3-1 or 3-2 (and certainly an attempt to shift to a four-dimensional construct wouldn't make things any easier). Thus it was necessary to simplify even further. Rather than actually trying to depict a continuous distribution on the four-dimensional simplex defined by the individual's remaining lifetime, we focused their attention on the central tendency of the relevant distribution. We then sought to convey the idea of a distribution using phrases such as "For each illness, we describe how it *might* affect you," "Consider the possibility that you might experience these two illnesses *around these times* in your life," or "starting *around* when you are 65 years old" [italics added for emphasis]

Obviously, describing each illness in terms of its expected latency period, its expected number of sick-years, recovered/remission years and lost life-years leaves each individual to infer the dispersion in the joint distribution of health states, since "around" and "about" remain unquantified. Our analysis therefore proceeds in terms of what we will assume respondents interpreted as the *expected values* of the distribution of future possible health profiles associated with the illness being described, as in Figure 3-4.

We hope future researchers can come up with sufficiently brief characterizations of the prospective joint distribution of multiple future health states to allow them to improve upon our approach to eliciting preference over risk reductions concerning a wide variety of future illness profiles.

3.7 Costs

Individuals' willingness to pay for risk reductions is expected to vary systematically with age. Very few young people are likely to be willing to pay large monthly costs to reduce risks that they do not feel will be relevant to them for many decades. In contrast, the same

risks will be highly salient to older respondents. Thus, we define three age brackets and draw program costs randomly from a different distribution for each group.

Eligible program costs by age group are shown in Table 3-4. (For the Canadian sample, the contemporaneous exchange rate placed one Canadian dollar at approximately $0.64 U.S.) It is useful to benchmark the plausibility of these program costs. In the case of blood test programs, a monthly cost of $3 translates into a fee for the annual test of $36, whereas monthly fees of $50, $90 and $140 translate into annual test costs of $600, $1080 and $1680. For the vehicle upgrades that prevent traffic accidents, these costs might be amortized over, say, 5 years, leading to a minimum equipment cost of $180 at the $3 monthly fee and costs of $50, $90 and $140 translate into equipment costs of $3,000, $5,400 and $8,400.

3.8 No strict dominance in risk reduction and cost

Combinations of risk reductions and costs are screened and rejected if there is "strict dominance" in the sense that a program that produces a greater risk reduction involves a smaller cost. Remaining cases are characterized by a risk-reduction/cost *tradeoff*, in the sense that programs that provide bigger reductions in risk more will always cost more. Given that we were concerned primarily about our prospective ability to identify any statistically significant relationship between choices and illness profiles and program costs, we were reluctant to dilute the design with choice scenarios that did not force the respondent to make tradeoffs between money and risk reductions. Allowing strict dominance would have created the opportunity for individuals to display aberrant choice behavior by choosing the non-dominant alternative. However, if this happens very infrequently, those choices are in a sense "wasted." Given that a substantial proportion of such apparently aberrant behavior is likely to be eliminated by excluding individuals who do not pass the risk comprehension test, we opted to exclude strict dominance in choices along these two key dimensions.

3.9 Disease latencies

The delay between now and the time of onset of the illness or injury is drawn from a uniform distribution on the integers between 1 and 60, subject to a number of rejections and redraws according to the credibility of different disease profiles. Randomly drawn latencies are compared to current age and to the respondent's hypothetical life expectancy and re-drawn if necessary. See Table 3-5.

3.10 Durations of illness/injury spells

Early phases of survey development distinguished four possible period within an illness or injury episode: leading months of moderate pain/disability (LEADMOD), months of severe pain/disability (SEVERE), trailing months of moderate pain/disability (TRAILMOD), and months between time of recovery or remission and death (TODEATH). We continued to generate lengths of spells within this framework, although we eventually aggregated the leading and trailing months of moderate pain/disability and presented to the respondent only the *total* number of months at each severity level in the survey instruments actually used to collect the data for these studies.

Eligible patterns of pain and disability (disease profiles) vary with the name of the disease. From these eligible patterns, we make random draws, which are then screened further to preclude implausible combinations. The eligible patterns are as shown below. Zero values indicate that there is no such spell in the particular profile. If TODEATH is zero, the person dies directly at the end of the last non-zero spell. Cases of sudden death have zero values for all four spells. Values of 999 signify that the spell is open-ended. Each value in the lists in Table 3-6 is equally likely, so repetitions of values signify that they are more likely than values that appear in the list only once. Following any particular spell that is open-ended, all subsequent spells will also be coded as open-ended, but the *first* open-ended spell determines the individual's state until death.

3.11 Randomization exclusions (i.e. redraw criteria)

The randomization strategy involved independent draws from all of the separate distributions described above, followed by rejections with completely new sets of draws under each of the following conditions:

a.) Require roughly equal representation of mixed and unmixed pain/disability levels.

- The randomly drawn illness/injury profiles are subjected to a number of screens before being accepted. In 50% of cases, we reject "mixed" profiles, defined as having more than one level of pain/disability. There will be some naturally occurring instances where SEVERE is 0 or both LEADMOD and TRAILMOD are both zero, and these cases will remain despite this screening, so at least 50% of cases will be characterized by only one level of pain/disability.

b.) Reject if start of pain spell is beyond life expectancy.

- This will be largely precluded by our separate screening of latencies

c.) Reject if life expectancy is in the middle of a closed-ended spell.

- Only open-ended pain/disability intervals should reach all the way to the individual's hypothetical life-expectancy

d.) Reject if closed-ended pain spell concludes, with either recovery or death, beyond life expectancy.

e.) Reject if it is a traumatic illness/injury (heart attack, stroke, traffic accident) and there is LEADMOD before SEVERE.

f.) Reject if they recover from the spell in question, but die from something else within a year

- Must live at least a year if they recover.

g.) Redraw if positive TRAILMOD, but no SEVERE (make it all LEADMOD).

h.) Redraw if positive TRAILMOD but no SEVERE or LEADMOD

- Make sure it is just LEADMOD in these types of cases.

i.) *Redraw if there is no LEADMOD, SEVERE, or TRAILMOD and you do not die immediately (TODEATH=0).*

- If there is no illness, it is not a "case" unless you are killed suddenly.

j.) *In mixed cases, with both moderate and severe pain/disability, if there is any SEVERE pain/disability followed by recovery, there must be at least one month of TRAILMOD*
   - Instant cures are not really plausible.

k.) *The first and second illness profiles cannot be the same*
   - Since these two profiles are used in the "training" phase of the survey instrument and it would be confusing to respondents if the disease profiles randomly ended up being identical.

l.) *It is possible to impose limits on the maximum duration of spells*
   - Either for all alternatives, or just for the first pair, but this limitation is not currently activated.

m.) *Additional adjustments*
   - A number of adjustments were made for several illness labels, to improve the realism of the choice sets. See Table 3-7.

3.12 Conversion to prose of the quantitative data, and rounding

Pre-tests of the survey instrument revealed that respondents were confused by too much spurious precision in the description of the age of onset, the lengths of pain/disability spells, and the age at death. Thus, the current version of the survey uses integer ages of onset, and rounds age at recovery or death to the whole year of age that the respondent will have attained at that time. Moderate and severe pain/disability spells are displayed in months up to 23 months and are rounded to whole numbers of years beyond that. Numbers actually displayed are the same numbers used in the analysis.

3.13 Arrangement of illness/injury spell data in choice tables

Within the survey instrument, the pain/disability descriptions are presented in a two-line pair. If the scenario is not mixed, so that there is only a single level of pain, then this pain description and its duration occupy the first line and the second line is blank. When there are two levels of pain, the moderate pain is presented first, with two exceptions:

i.) if the moderate pain is open-ended ("for remaining life") and the severe pain is finite, then the moderate pain is described second;

ii.) if the individual recovers from the pain spell (as opposed to the case being terminal), the severe pain duration is presented first and the moderate pain phase second, to imply a recuperation period.

It may have been ideal to preserve all three pain spells which were employed in our earlier survey variants that used visual profiles. With a three-line description, we could

have preserved the time profile of moderate and severe illness, but we decided that this created more complexity than was warranted.

3.14 Hospitalization

Each illness/injury is also characterized by the duration of conventional hospitalization it entails. We assume that this does not include single-day emergency room events. Three days is the shortest period of hospitalization in our design.

Durations for a specific illness are constrained by the duration of that illness, but are otherwise drawn from a set of eligible durations for each illness. If the "draw" exceeds the duration of the illness, another draw is made.

In the case of Alzheimer's disease, the hospitalization is described differently (i.e., as "long-term care").

Our algorithm rejects cases where the individual is hospitalized for longer than they are sick, or hospitalized when there is sudden death (described as within a few hours). We also reject cases where there is severe pain from traumatic illness/injury but no hospitalization.

Eligible durations of hospitalization are given in Table 3-8. When the outcome is "die suddenly", there is no hospitalization.

3.15 Surgery

Each illness/injury is also characterized by the type of surgery that would be involved, if any. The table describes eligible possibilities. As draws are made, our algorithm rejects cases that are ineligible.

Eligible descriptions of surgery are shown in Table 3-9. When the outcome is "die suddenly", there is no surgery.

3.16 Orthogonality

In our main paper, there are several key variables which we use to explain individuals' choices among risk-reduction programs. In the estimating specification, these constructed variables are first subjected to nonlinear monotonic transformations, to allow for diminishing marginal utility. The basic variables are: income, Y, and program cost, c (combined to yield the net income variable); the present discounted time spent in each of three adverse health states: illness-years, $pdvi$, recovered/remission years, $pdvr$, and lost life-years, $pdvl$; and the respondent's age.[10]

Other than the income and age variables, the illness profiles and program costs are randomly assigned. However, these random assignments are subject to the requirement that nobody should consider an illness profile where they get sick at an age younger than their current age, and that none of the illness profiles should be implausible, such as sudden death, with no illness from diabetes or Alzheimer's disease, or recovery from either of these conditions).

[10] Of course, other relevant variables are the size of the risk reduction, which is assigned randomly, and in models which are non-linear in net income, the size of the baseline risk, also assigned randomly.

Whenever some combinations of attributes are precluded on plausibility grounds, it is important to ask whether the remaining explanatory variables remain sufficiently orthogonal to permit the estimation of slope coefficients without too many problems with multicollinearity. Table 3-10 gives the correlations between the simplest forms of the main explanatory variables in our models.

It is not particularly surprising that the age variable should be somewhat negatively correlated with the other variables, which are present discounted quantities. The older a respondent is, the fewer will be the remaining life years over which they can experience spells of latency, sick-time, recovered/remission time, or lost life-years.

Plots of the relationships between the net income variable and age, and between the net income variable and the illness-years variable, are provided in Figure 3-5 and Figure 3-6. While the variance of the net income term is less than that for the other two variables, there appears to remain plenty of independent variation in these modestly correlated variables.

3.17 Section 3 Tables

Table 3-1 Complete joint distribution of age and gender
in our estimating sample (n = 1,801)

age	# male	# female	Total	Age	# male	# female	Total
25	7	14	21	59	18	15	33
26	15	18	33	60	20	23	43
27	9	13	22	61	13	18	31
28	23	9	32	62	11	23	34
29	20	24	44	63	19	10	29
30	16	12	28	64	14	18	32
31	15	16	31	65	11	17	28
32	18	9	27	66	11	14	25
33	15	17	32	67	8	17	25
34	17	14	31	68	13	16	29
35	14	10	24	69	7	11	18
36	18	14	32	70	13	17	30
37	12	18	30	71	9	16	25
38	20	22	42	72	12	20	32
39	23	27	50	73	9	12	21
40	18	17	35	74	6	19	25
41	27	20	47	75	7	10	17
42	21	19	40	76	6	8	14
43	27	23	50	77	9	9	18
44	37	23	60	78	7	8	15
45	26	26	52	79	6	3	9
46	18	21	39	80	2	5	7
47	19	17	36	81	4	4	8
48	18	24	42	82	4	5	9
49	19	21	40	83	5	3	8
50	20	17	37	84	2	5	7
51	23	24	47	85	3	0	3
52	18	14	32	86	1	1	2
53	16	17	33	87	0	1	1
54	11	19	30	88	1	2	3
55	21	18	39	90	1	1	2
56	12	19	31	93	0	2	2
57	7	13	20				
58	7	20	27	**Total**	**859**	**942**	**1,801**

Table 3-2 Range of attributes used for illness profiles

By illness name; means and standard deviations; estimating sample = 1,801 different individuals, 7,520 completed choice occasions, 15,040 illness profiles, 22560 alternatives. See Section 3.3.

Health Threat:	1 Breast Cancer (females)	2 Prostate Cancer (males)	3 Colon Cancer	4 Lung Cancer	5 Skin Cancer	6 Heart Attack
# profiles	697	676	1357	1368	1353	1406
Monthly cost ($)	30.78 (30.09)	28.12 (26.09)	29.35 (28.37)	30.4 (28.7)	30.19 (28.81)	29.85 (29.62)
Risk reduction	0.0033 (0.0016)	0.0034 (0.0017)	0.0034 (0.0017)	0.0034 (0.0017)	0.0035 (0.0017)	0.0035 (0.0017)
Latency (pre-illness years)	17.0 (11.0)	18.5 (11.2)	18.4 (11.6)	19.4 (11.5)	17.6 (11.7)	20.5 (12.5)
Illness years	4.9 (3.5)	4.9 (3.9)	8.5 (8.3)	8.3 (7.7)	7.5 (7.3)	3.4 (6.6)
Lost life-years	11.5 (11.4)	12.0 (11.5)	8.9 (9.7)	10.3 (9.8)	10.3 (10.8)	13.5 (11.3)
1(sudden death)	0	0	0	0	0	0.52
1(recovery/ remission)	0.60	0.64	0.39	0.23	0.40	0.19
Health Threat:	7 Heart Disease	8 Stroke	9 Resp. Disease	10 Traffic Accident	11 Diabetes	12 Alzheim. disease
# profiles	1423	1424	1337	1295	1357	1347
Monthly cost ($)	29.87 (28.63)	30.85 (29.43)	29.77 (29.41)	29.72 (27.92)	29.17 (28.07)	29.84 (28.54)
Risk reduction	0.0034 (0.0017)	0.0034 (0.0017)	0.0034 (0.0017)	0.0034 (0.0017)	0.0033 (0.0016	0.0033 (0.0016)
Latency (pre-illness years)	19.4 (11.9)	21.8 (12.7)	21.4 (12.2)	18.2 (12.3)	18.2 (10.8)	22.6 (12.5)
Illness years	10.2 (8.8)	3.6 (6.4)	7.4 (6.5)	4.0 (7.6)	6.8 (5.8)	6.8 (4.7)
Lost life-years	7.4 (8.4)	12.0 (10.1)	8.0 (7.8)	14.5 (12.5)	13.4 (10.7)	8.8 (6.4)
1(sudden death)	0	0.51	0	0.51	0	0
1(recovery/ remission)	0.26	0.19	0.38	0.19	0	0

Table 3-3 Eligible risks per 1000 and percent risk reductions

(See Section 3.5)

Baseline risk	Percentage Reductions
4	25,50,75
5	20,40,60,80
8	25,50,75
10	10,20,30,40,50,60
15	20,40
20	5,10,20,30
30	10,20
40	5,10

Table 3-4 Eligible program costs by age group

(See Section 3.7)

Age group	**Program costs (per month)**
Under 40	2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 16; 18; 20; 25; 50
40-65	4; 6; 8; 10; 12; 13; 14; 15; 16; 17; 18; 19; 20; 25; 30; 35; 40; 45; 50; 55; 60; 65; 70; 75
Over 65	5; 10; 15; 20; 25; 30; 35; 40; 45; 50; 55; 60; 65; 70; 75; 80; 85; 90; 95; 100; 105; 110; 115; 120; 125; 130; 135; 140

Table 3-5 Latency exclusions

(See Section 3.9)

Illness/Injury	Re-draw random latency if:
Gender-specific Cancer	latency<3 or age+latency<30 or age+latency>=life expectancy
Colon Cancer	latency<3 or age+latency>=life expectancy
Lung Cancer	(latency<5 and age<65) or (latency<2 and age>=65) or age+latency<40 or age+latency>=life expectancy
Major Skin Cancer	latency<2 or age+latency>=life expectancy
Heart Attack	latency<3 or age+latency<40 or age+latency>=life expectancy
Heart Disease	latency<3 or age+latency<40 or age+latency>=life expectancy
Stroke	latency<3 or age+latency<50 or age+latency>=life expectancy
Lung Disease	(latency<5 and age<65) or (latency<2 and age>=65) or age+latency<50 or age+latency>=life expectancy
Traffic Accident	latency<1 or age+latency>=life expectancy
Diabetes	(latency<5 and age<65) or (latency<2 and age>=65) or age+latency>=life expectancy
Alzheimer's Disease	(latency<5 and age<65) or (latency<2 and age>=65) or age+latency<60 or age+latency>=life expectancy

Table 3-6 Eligible durations within illness/injury profiles

(Values of 999 signify that the spell is open-ended, bounded by whatever is the individual's nominal life expectancy, which depends on their current age; note that LEADMOD and TRAILMOD were ultimately combined and the exact timing of moderate and severe spells of pain and disability within the period of sick-years was left unspecified. See Section 3.13Arrangement of illness/injury spell data in choice tables**)**

Spell type	Eligible durations (months)
Gender-specific Cancers	
- LEADMOD	1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60
- SEVERE	0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60
- TRAILMOD	0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60
- TODEATH	0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 1; 2; 3; 6; 9; 12; 18; 24; 30; 36; 42; 48; 54; 60; 999; 999; 999; 999; 999; 999; 999; 999; 999; 999
Colon Cancer	
- LEADMOD	1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- SEVERE	0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TRAILMOD	0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TODEATH	0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 1; 2; 3; 6; 9; 12; 18; 24; 30; 36; 42; 48; 54; 60; 999; 999; 999; 999; 999; 999; 999; 999; 999; 999
Lung Cancer	
- LEADMOD	1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- SEVERE	0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TRAILMOD	0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TODEATH	0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 1; 2; 3; 6; 9; 12; 18; 24; 30; 36; 42; 48; 54; 60; 999
Serious Skin Cancer	
- LEADMOD	1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- SEVERE	0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60
- TRAILMOD	0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TODEATH	0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 1; 2; 3; 6; 9; 12; 18; 24; 30; 36; 42; 48; 54; 60; 999; 999; 999; 999; 999; 999; 999; 999
Heart Attack	

- LEADMOD 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- SEVERE 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TRAILMOD 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TODEATH 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999;999;999

Heart Disease

- LEADMOD 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999;999;999
- SEVERE 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999;999;999
- TRAILMOD 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999;999;999
- TODEATH 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999;999;999

Stroke

- LEADMOD 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- SEVERE 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TRAILMOD 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TODEATH 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999;999;999

Respiratory Disease

- LEADMOD 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- SEVERE 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TRAILMOD 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TODEATH 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999;999;999

Traffic Accident

- LEADMOD 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- SEVERE 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TRAILMOD 0; 0; 0; 0; 0; 1; 2; 3; 6; 9;12;18;24;30;36;42;48;54;60;999;999
- TODEATH 0; 0; 0; 0; 0; 1; 2; 3; 6; 9; 12; 18; 24; 30; 36; 42; 48; 54; 60; 999; 999; 999; 999; 999; 999; 999; 999; 999; 999

Diabetes (under 65 respondent)

- LEADMOD 36;48;60;72;84;96
- SEVERE 0; 3; 6; 9; 12; 18; 24; 30; 36; 42; 48; 54; 60; 999; 999; 999; 999
- TRAILMOD 0
- TODEATH 0

Diabetes (65 or over respondent)

- LEADMOD 12;24;36;48;60;72;84;96
- SEVERE 0; 3; 6; 9; 12; 18; 24; 30; 36; 42; 48; 54; 60; 999; 999; 999; 999
- TRAILMOD 0
- TODEATH 0

Alzheimer's Disease (under 65 respondent)

- LEADMOD 36;48;60;72;84;96
- SEVERE 0;12;24;36;48;60;72;84;96;999;999;999;999
- TRAILMOD 0
- TODEATH 0

Alzheimer's Disease (65 or over respondent)

- LEADMOD 24;36;48;60;72;84;96
- SEVERE 0; 6; 9; 12; 18; 24;30;36;42;48;54;60;72;999;999;999;999
- TRAILMOD 0
- TODEATH 0

Table 3-7 Illness/injury profile adjustments

(See Section)

Illness/Injury	Adjustments
Gender-specific Cancer	none
Colon Cancer	none
Lung Cancer	none
Major Skin Cancer	none
Heart Attack	Select 15% randomly and convert to “die suddenly” (all spells zero)
Heart Disease	none
Stroke	Select 15% randomly and convert to “die suddenly” (all spells zero)
Lung Disease	none
Traffic Accident	Select 15% randomly and convert to “die suddenly” (all spells zero)
Diabetes	If age<65, no severe spell unless moderate spell for at least 60 months; If age>=65, no severe spell unless moderate spell for at least 36 months
Alzheimer’s Disease	If age<65, no severe spell unless moderate spell for at least 60 months; If age>=65, no severe spell unless moderate spell for at least 36 months

Table 3-8 Eligible durations of hospitalization

Values of 999 to be interpreted as "for remainder of nominal lifespan" which will depend upon the respondent's current age.

Illness/Injury	List from which hospitalization period is randomly drawn (expressed in months or fractions thereof; 999=open-ended)	Comments/Assumptions
Gender-specific Cancer	0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0	at least some hospitalization
Colon Cancer	0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0	at least some hospitalization
Lung Cancer	0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0	at least some hospitalization
Major Skin Cancer	0; 0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0	hospitalization not required
Heart Attack	0; 0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0	hospitalization not required
Heart Disease	0; 0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0	hospitalization not required
Stroke	0; 0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0; 999.0; 999.0; 999.0; 999.0; 999.0; 999.0; 999.0	hospitalization not required
Lung Disease	0; 0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0	hospitalization not required
Traffic Accident	0; 0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0; 999.0; 999.0	hospitalization not required
Diabetes	0; 0.10; 0.25; 0.50; 1.0; 1.5; 2.0; 2.5; 3.0; 4.0; 5.0; 6.0; 999.0	hospitalization not required
Alzheimer's Disease	3; 6; 12; 24; 36; 48; 60; 72; 84; 96; 999.0; 999.0; 999.0; 999.0; 999.0; 999.0; 999.0; 999.0	Cannot escape eventual long-term hospitalization, but this is cast as "long-term care"

Table 3-9 Eligible surgery descriptions
(See section)

Illness/Injury	Surgery Descriptions	Comments/Assumptions
Gender-specific Cancer	None; minor; major	none
Colon Cancer	Minor; major	none
Lung Cancer	None; minor; major	none
Major Skin Cancer	Minor; major	none
Heart Attack	None; minor; major	none
Heart Disease	None; minor; major	none
Stroke	None; minor; major	none
Lung Disease	None; minor; major	none
Traffic Accident	None; minor; major	none
Diabetes	None	none
Alzheimer's Disease	None	none

Table 3-10 Correlations among estimating variables
(See Section)

Variable:	1	2	3	4	5
1 $\left(Y_i - c_i^j\right)^{(0.42)} cterm_i^j - \left(Y_i\right)^{(0.42)} yterm_i^j$	1				
2 $\Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$	0.2341	1			
3 $\Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	0.0807	0.1214	1		
4 $\Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	0.0487	0.1465	0.0382	1	
5 age	-0.369	-0.1191	-0.0137	-0.0489	1

For n = 15040 illness-profile/risk-reduction combinations, where all variables except age are zero for "No Program" alternative

3.18 Section 3 Figures

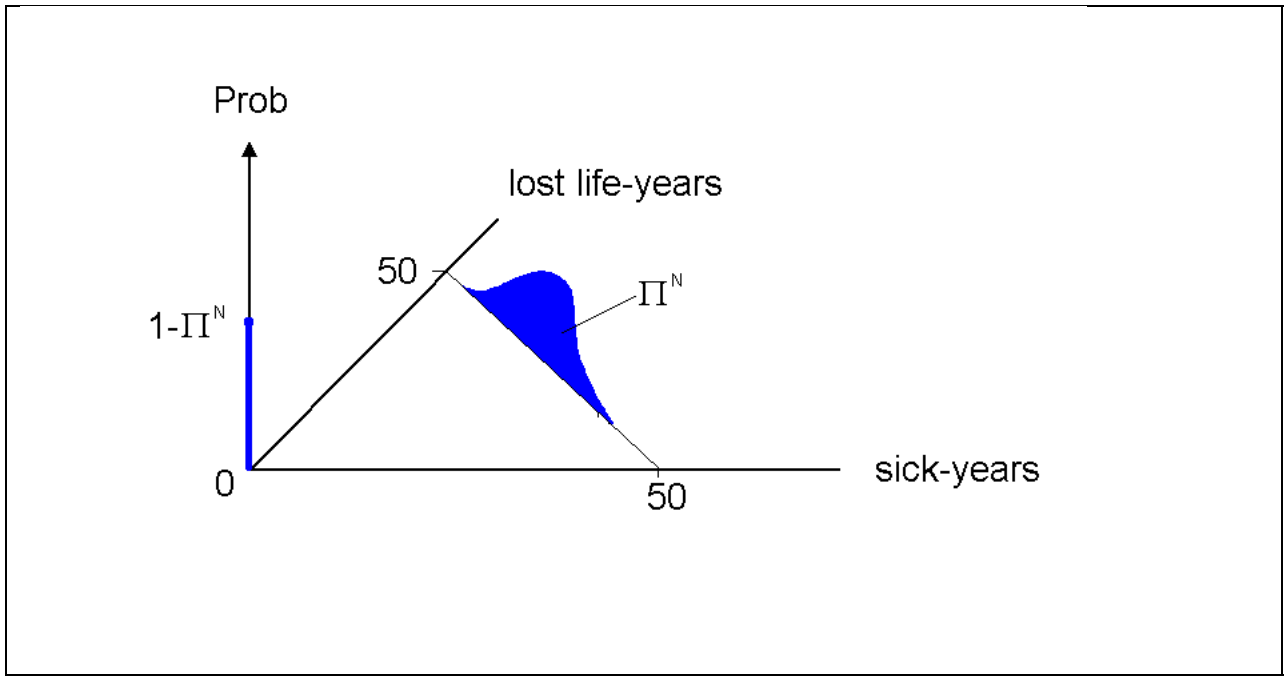


Figure 3-1 Distribution of health profiles without risk-reduction program

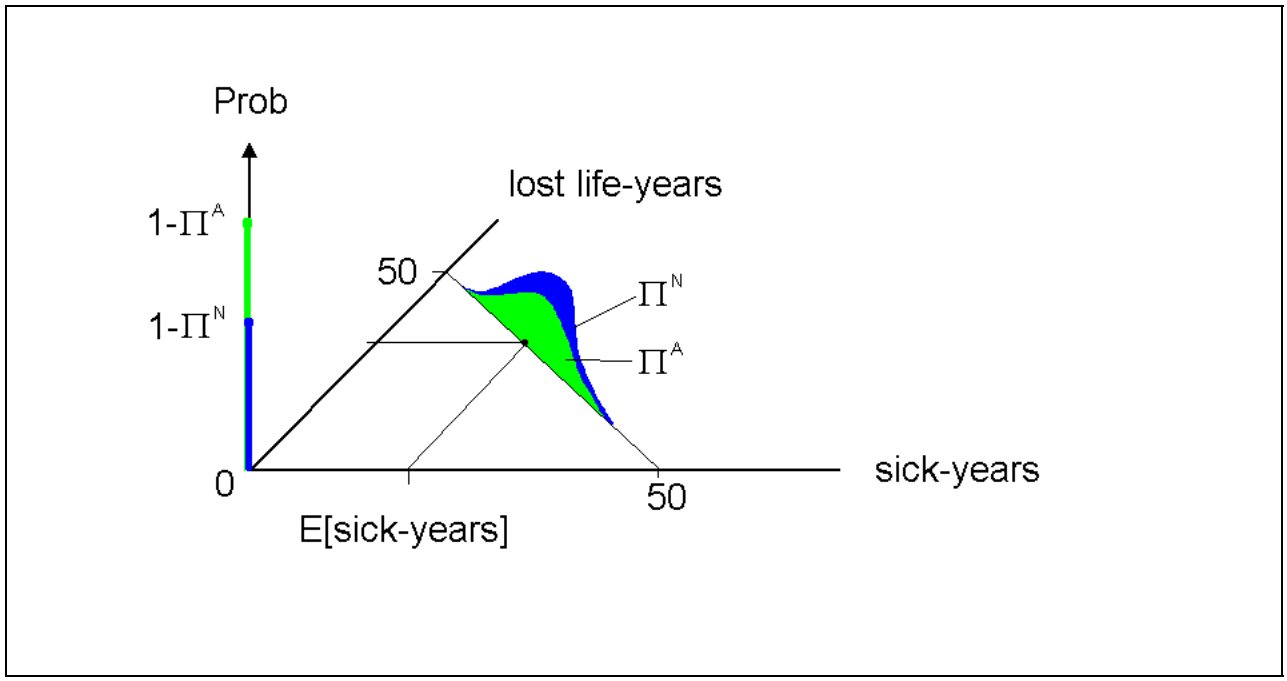


Figure 3-2 If program changes probabilities but not illness profiles

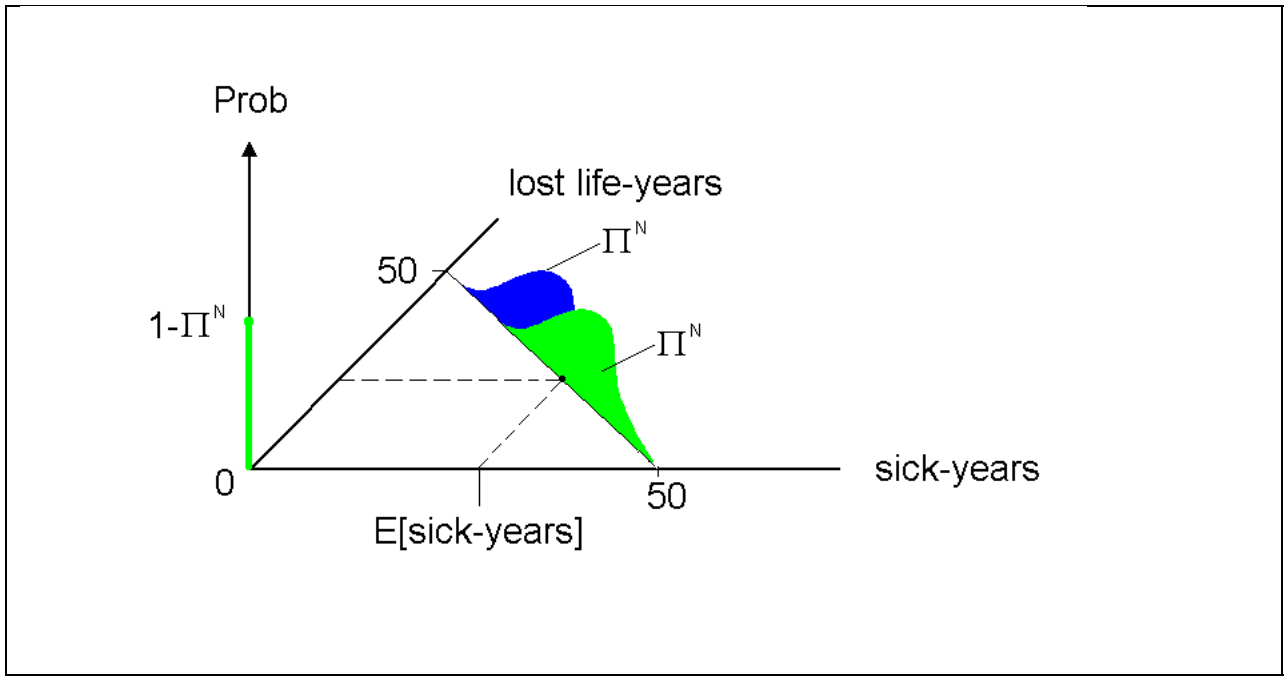


Figure 3-3 If program changed illness profiles but not probabilities

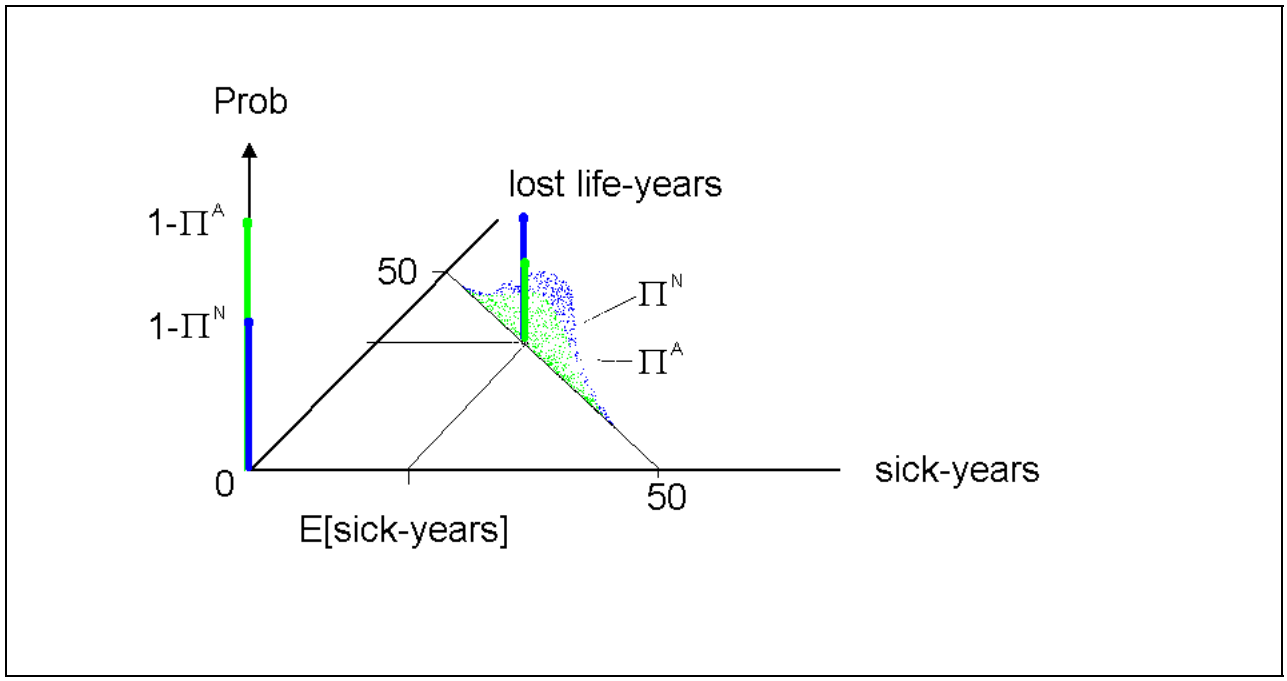


Figure 3-4 Illness profile involves "about" 20 sick-years, 30 lost life-years

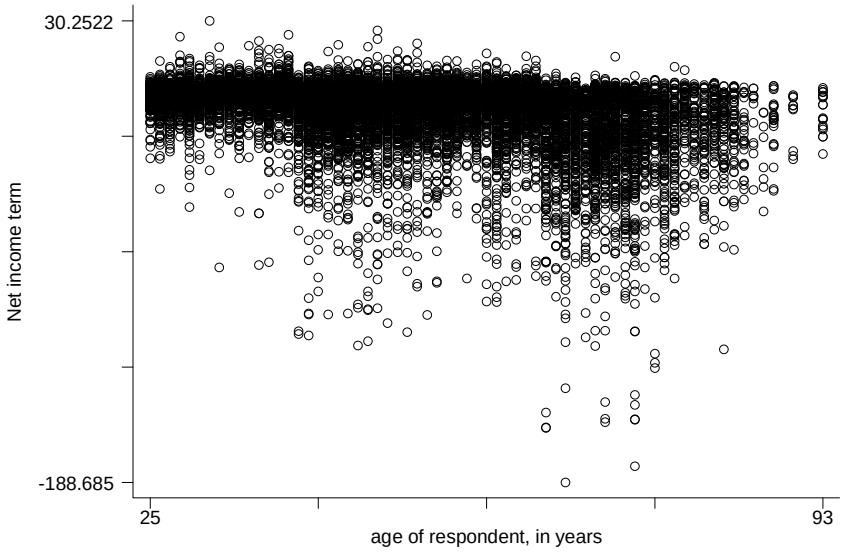


Figure 3-5 Joint distribution of net income terms and age of respondent

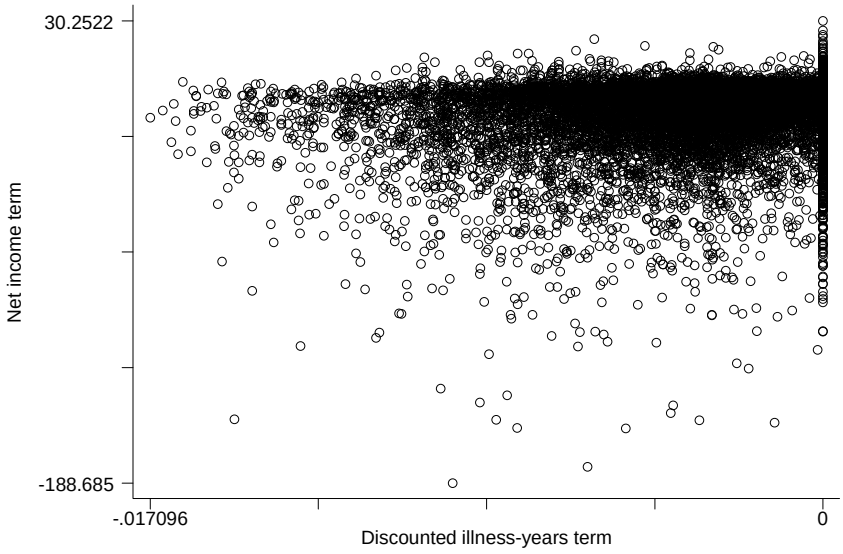


Figure 3-6 Joint distribution of net income term and discounted illness-years term

4 The Knowledge Networks Panel and Sample Selection Corrections

4.1 Introduction

Equation Section (Next)

Our central concern was achieving the highest quality and most representative sample possible since we wish to be able to extrapolate our findings to the U.S. population as a whole. This required (1) selecting the best survey firm based on the quality of its sample, (2) excluding observations based only upon the most conservative ex ante criteria, and (3) evaluating and correcting to the extent possible for any sample selection bias that has occurred. In this section, we describe how we accomplished each of these tasks.

In recent years, online survey methods have made rapid gains in popularity among researchers. Deutskens et al. (2006) note that by 2004 about 35% of the U.S. survey research market consisted of online surveys.[11] A large number of survey research firms now offer this mode of delivery (see Evans and Mathur (2005) and Wright (2005)). The online survey mode is attractive because it allows researchers to reduce field costs and improve response and data processing times. Despite these advantages, the sampling properties of these surveys can be less than ideal. As Best et al. (2001) note, most Internet sampling procedures "only permit the generation of diverse, not representative, samples." Much recent effort has been devoted to assessing the representativeness of online surveys as compared to traditional random-digit-dialed (RDD) telephone surveys or mail surveys.[12]

4.2 Survey firm qualifications and sample properties

Our decision about which firm to select was based on the quality of the sample that could be offered. Two of the leading U.S. survey research firms at the time were Knowledge Networks, Inc. (KN, formerly Intersurvey) and Harris Interactive, Inc. (HI, formerly Harris Black International).[13] There are a variety of ways to recruit members for an online survey panel.[14] At the time of our survey, Knowledge Networks recruited its panelists via an initial RDD contact and equiped panelists who did not have computers or internet access with Web-TV hardware and internet access. Knowledge Networks relied upon its RDD recruiting methods to ensure a maximally representative panel.[15] For this reason, we chose the KN panel for our study.[16] In

[11] See discussion of the state-of-the-science in survey research in Tourangeau (2004). A number of relevant concerns are also outlined in Birnbaum (2004).

[12] Ilieva et al. (2002), Schonlau (2004), Schillewaert and Meulemeester (2005) address the (relative) sampling properties of web-based or email surveys. Web-based panels, in particular, have been addressed by Lee (2006). Some social science disciplines, such as economics, have struggled with sample selection bias detection and correction for decades (e.g. going back to early work by Heckman (1979), with the broader scope of the early work surveyed by Vella (1998)). Winship and Mare (1992) summarized the issue for sociologists. In other social science disciplines, these issues have been addressed routinely only in more recent years (e.g., Cuddeback et al. (2004) describe the state of practice in social work research).

[13] See http://www.knowledgenetworks.com and http://www.harrisinteractive.com.

[14] Some of the possibilities have been explored systematically by Goritz (2004), for example.

[15] Smith (2003) compares answers from the General Social Survey (GSS) with answers to the same questions by a sample of KN panel members. They find many similarities, but a few differences.

contrast, Harris uses a wide range of recruitment methods, but panel membership is conditional on the panelist already having web access capability. Berrens et al. (2003) provide a description of the recruitment methods used by each firm at the time our survey was being developed.[17]

Harris Interactive acknowledges that its recruitment methods do not yield a representative panel, but it has developed a method using "propensity scores" to construct post-stratification weights to adjust the relative influence of different panelists.[18] These propensity scores are based on an array of benchmark attitudinal questions posed both in each online survey and in Harris' regular RDD "reference surveys." Berrens et al. (2003), Schonlau et al. (2004), and Duffy et al. (2005) describe how HI pools the data on these attitudinal questions across an online survey and their most current reference survey, using an indicator for the source of the data as the dependent variable in a logistic regression. The fitted values for the systematic portion of this regression (the propensity scores, or the associated conditional probabilities) are sorted into quintile or decile bins. These bins constitute an additional dimension (along with a number of study-specific observable sample characteristics such as race, gender, age, and income that may be used separately, or as part of the same logistic regression) to construct weights for each online survey observation that render its influence comparable to the influence of the same category of individual in the general population.[19]

4.3 Estimating sample

4.3.1 Comparison to 2000 Census distributions of age, income, gender

Table 4-1 describes the marginal distributions of the estimating sample of 1,801 subjects against the distribution in the 2000 Census. While we requested that panels 25 years and older should be invited to participate, the sample includes a few people who are still 24 years old. These individuals are included in the 25-34 year old group. Notice that as a percentage of the 25-year-and-older population, our sample has slightly fewer young people and slightly more older people, except in the 75-year-and-older group. In the face of priors that might suggest that an internet- and Web-TV-based survey might select systematically in favor of younger people, this is reassuring.

Given the difficulty in eliciting accurate information about overall household income, we are also very satisfied with the correspondence between the distribution of income brackets in our estimating sample and the corresponding distribution in the 2000 Census. The gender mix is also representative of the population.

---

[16] Respondents were also paid ten dollars for completing our survey, in addition to the usual benefits of Knowledge Networks panel membership. More information about Knowledge Networks is available from their website: www.knowlegdenetworks.com.

[17] In reaction to concerns about validity of the inferences from online surveys, the *Journal of Medical Internet Research* has proposed a checklist of recommendations for authors in an effort to ensure complete descriptions of Web surveys (Eysenbach (2004)).

[18] Simple post-stratification weights, based upon the relative frequencies of types of respondents in the sample versus the population (say, according to a recent census data) have been discussed in many studies. The viability of this strategy has been assessed for email-based surveys by Best and Krueger (2002), and for web-based surveys by Bandilla et al. (2003).

[19] A similar technique, based on method of moments estimation, has been demonstrated by Nevo (2003).

4.3.2 Exclusion criteria for our estimating sample

For the panelists who responded to our survey, a total of 11,717 risk-reduction program choice questions were answered. We used just three conservative a priori criteria to disqualify choice sets from the estimating sample.

4.3.2.1 A minor choice set design error (2.8% of the sample)

First, in a number of choice scenarios, our algorithm that assigned latency time, sick-time, recovered/remission time and lost life-years produced a projected age at death that exceeded the individual's nominal life expectancy. This was unintended, and potentially implausible for most of these major illnesses, so we excluded all choice scenarios containing an illness profile with this feature. This criterion affected 331 of our 11,717 choices.

4.3.2.2 Inadequate risk comprehension (a further 18% of the sample)

The second reason to exclude choices from the estimating sample was for individuals who could not correctly answer the "risk comprehension" question on our survey. (Form numbers referenced in the following pertain to the single example of our surveys provided at the end of this document. Note: special quotes, «», denote fields which were tailored to the individual and/or randomized across choice sets. The quotes themselves do not appear in the actual choice sets.) Figure 4.1 shows the key risk information and the risk comprehension question presented on **Form 20.**

Of course, it is possible that some people did in fact understand the risk information in the survey, but merely answered this particular question incorrectly, by mistakenly clicking the wrong button. These people's choices will have been disqualified unnecessarily. Likewise, it is possible that some people did not understand the risk information at all, but randomly chose a button for this question and got the answer right. These people's choices will have been included in our sample despite these individuals having a poor understanding of the risk information provided in the choice sets. There is no way of knowing the extent to which an occasional random guess by a respondent has led us (a) to exclude people who *do* understand risk, or (b) to include those who *do not* understand risk. However, we are more concerned about the second type of error, so we exclude 2120 additional choices from our estimating sample (equivalent to five choices for each of 424 respondents, and 18% of the original 11,717 choices).

4.3.2.3 Outright scenario rejection (a further 15% of the sample)

The third reason to exclude choices from the estimating sample is an "outright scenario rejection" criterion. We assume that if an individual chose one of the two offered programs that they were "playing along" with the stated preference choice exercise and were willing to incur the annual cost of a program in exchange for the risk reduction that is described for that alternative. However, when an individual chooses "Neither Program" it is not clear whether they were "playing along" with the stated preference choice exercise and found both programs too costly for the benefits that they perceive, or whether they were rejecting the choice scenario as implausible. Following each choice set, therefore, we invited those respondents who chose the "Neither Program" alternative to elaborate upon their reasons for doing so. Included among the list of reasons was the set of reasons suggested in the earlier "cheap talk" script on **Form 22**, which is reproduced as Figure 4-2. Each of these suggested reasons is a legitimate "economic"

reason why an individual might choose "Neither Program" —a reflection of the individual's budget constraint or their preferences.

However, we seek to identify those cases where the individual's choice of the status quo option, "Neither Program," reflects a failure to play along with the stated choice exercise. In only those cases where the respondent selected "Neither Program," they were asked the question appearing on **Forms 28, 32, 36, 40,** and **44**, reproduced in Error! Reference source not found.**.** Some respondents selected "Neither Program" and indicated, in this follow-up question, that the ONLY reason why they did not want to pay was "I did not believe these programs would reduce my risks." Prior to the choice exercises, they had been given specific instructions on **Form 24** to assume that the programs would work as advertised, as reproduced in Figure 4-4. These respondents, therefore, were not "playing the game" as they had been instructed. If a respondent listed any of the *other* reasons offered in the list, they were given the benefit of the doubt and retained in the sample, since it is not possible to tell whether the economic reasons for choosing "Neither Program," or their doubts about the program's effectiveness, had dominated.[20]

This third criterion for dropping choices from the estimating sample affects 1745 additional choices, about 15% of the original 11,717, and the equivalent of five choices for each of 349 respondents (although not all individuals rejected all five of their choice sets).

4.3.2.4 Other possible exclusion criteria (not implemented)

The remaining sample is used in our estimating specifications. It consists of a total of 22,560 alternatives contained in 7,520 choice sets considered by 1,801 different individuals (although not every individual provides the full set of five choices).

As due diligence, however, we also explored the consequences of further exclusions based upon arbitrary criteria for how much time respondents spent, in total, on their five different choice exercises. The loosest criterion was to drop respondents if they did not spend at least 60 total seconds (average 12 seconds per choice set) on the choice tasks (47 additional people), at least 80 seconds (average 16 seconds per choice set, for 55 additional people), or at least 100 seconds (average 20 seconds per choice set, for 77 additional people). Since the estimated parameters for our main specification were minimally affected by these further arbitrary exclusions based on response times, we did not implement these last three types of exclusions. People often make hasty decisions in revealed preference contexts as well, and we do not generally invalidate their choice information.

4.3.2.5 Effects on estimated parameters

Table 4-2 provides the results for one reasonably complete version our main estimating specification as each successive exclusion criterion described above is implemented. The results in column (4) in the table are the results for our final estimating sample with 22,560 alternatives (7,520 choices) made by 1,801 different individuals (although these models do not include the further sample selection correction or scenario adjustment variables discussed later in this document, so the estimates are slightly different from those featured in the main paper). It is clear from this table that exclusions for failure to correctly answer the risk comprehension question, and exclusions for outright scenario rejection concerning program effectiveness make

[20] In real health care decisions, of course, people will also decline to participate in health testing programs because they do not believe that these programs will reduce their risks.

the biggest difference to the estimated parameters. We are confident, however, that these two types of exclusions are imperative on a priori grounds.

4.4 OMB data quality standards

The U.S. Office of Management and Budget has recommended data quality standards for survey research when that research is intended to be used as the basis for policy decisions.[21] One specific dimension of these standards concerns the representativeness of survey samples. Earlier generations of researchers typically resorted to a simple assessment of the representativeness of such samples, consisting of side-by-side comparisons of the marginal distributions of key variables (such as age, income, and gender) for both the estimating sample and the relevant population. It is often straightforward to draw a sample in a manner that will ensure that the sample more or less matches the intended population in terms of the marginal distributions, and even the joint distribution, of the observable variables.

However, as most survey researchers now appreciate, there is a more subtle challenge. A sample that mimics the population in terms of the marginal distributions of a few *observable* variables may still be non-representative if the sample and the population differ in terms of *unmeasured or unobserved* characteristics. Correction methods such as the weights based on propensity score quantiles (as used by Harris Interactive) rely entirely on observed characteristics.[22] The effect of unobserved characteristics is especially relevant when the subject matter of the survey is more salient to some contacted households and less salient to others. Not all households will be equally inclined to participate in the survey. Furthermore, when using a standing consumer panel for survey research, it is not sufficient merely to compare those panelists who were invited to participate in a particular survey with those who actually chose to participate. The standing panel *itself* may be self-selected. One needs to reach all the way back to the random-digit-dialed recruiting contacts to assess representativeness.

In this section, we assess the potential for sample selection bias in the results from our survey sample drawn from the consumer panel maintained by Knowledge Networks, Inc. (KN). Our research goal is to determine whether representativeness appears to be adequately maintained—through the attrition, selection, and response process—so that our model, based on our estimating sample of 1,801 respondents, can be safely assumed to produce inferences that can be considered valid for the entire U.S. population.

[21] The Data Quality Act amends the Paperwork Reduction Act (44 U.S.C. 3501 et seq.) The DQA was enacted in December 2000 as a two-paragraph provision within an appropriations bill (see the Treasury and General Government Appropriation Act for Fiscal Year 2001, Pub. L. No. 106-554, § 515 Appendix C, 114 Stat. 2763A-153 (2000)). The DQA went into effect on October 1, 2002, which was a deadline for federal agencies to issue their final information quality guidelines. It is intended to apply to "influential scientific, financial, or statistical information," consisting of any data that will have an impact on significant public policies or major private sector decisions.

[22] Schonlau et al. (2004) acknowledge this limitation: “Propensity scoring balances observed covariates. Propensity scoring balances unobserved covariates only to the extent that they are correlated with observed covariates. The assumption that unobserved variables can be ignored with respect to selection bias is called ignorability.” These authors also concede that their weighting scheme adjusts their California sample to match the national distribution for the attitudes in the reference survey, but that “the additional assumption that the California population answers attitudinal questions just like the U.S. population…is not verifiable.”

4.5 Construction of selection model variables

Sample selection correction algorithms generally require that the researcher know something about each member of the population of interest that might help explain whether each individual appears in the final estimating sample. For this work, the "population" in question is the overall U.S. population. From that population, the random digit dialing algorithm should in principle produce a random subsample.[23] However, with random digit dialing, this means that the only thing one truly knows about every RDD residential contact attempt is the telephone number itself. Ideally, we would like to have individual-specific data on a wide variety of characteristics for the target of *every* RDD residential contact attempt, but this is impossible. Therefore, we use proxy data in the form of neighborhood characteristics at the census tract level by linking census tract data from the 2000 census to each household in the original KN panel recruitment sample frame.

The KN panel recruitment sample frame at the time of our survey included all working residential RDD phone numbers that KN first sampled and called (using the proprietary MSG Genesys-ID sampling system). While recruitment at KN is ongoing, the relevant recruiting phone numbers for our particular study sample were dialed between 1999 (when panel recruitment began) and May 1, 2002 (the date when the particular survey samples to be investigated were drawn for this health risk study). KN retained for analysis all valid residential phone numbers which included all cases with a final recruitment disposition code of "answering machine," "call back," "interview," "no answer," "refusal," and "refusal - privacy manager." The only exclusions from the original RDD sample were phone numbers found to be non-residential or non-working. These phone numbers are excluded because they are not explicitly associated with residential households. This recruiting strategy leaves roughly 525,000 unique phone numbers in the sample frame.

4.5.1 Linking KN RDD recruiting contacts to 2000 Census tracts

Of these over half-million phone numbers, roughly 400,000 had corresponding street addresses on file in the KN database (call this Subset 1). Some of these addresses came from reverse-address matching of just the phone numbers themselves, and others stemmed from telephone-based recruitment, where a telephone voice contact resulted in the contacted party providing a street address. Of these cases, about 80% had valid street addresses that could be successfully matched by ESRI's ArcView 3.3 and the ESRI StreetMap 2000 utility. These addresses were geocoded to identify approximate point locations (side of street and how far along block) for each residence. The approximate point locations of these residences were then overlaid with ESRI's census tract polygons, a standard GIS "theme" that is accompanied by an attribute file containing corresponding census tract data from the 2000 Census.

Of the remainder of the RDD telephone numbers with street addresses that could not be specifically matched by the StreetMap utility, most had usable zip code data (call this Subset 2).

[23] We must assume that households without telephone numbers are a sufficiently tiny fraction of the population that they can be ignored for most purposes. For studies targeting certain specialized groups, of course, this underlying selection problem could not be ignored (e.g., studies concerning the homeless). Since the early 2000s, of course, cell phone utilization has proliferated, resulting in multiple telephone numbers per household. Knowledge Networks now recruits panelists based on address sampling, rather than randomly dialed telephone numbers.

These cases were matched, albeit less accurately, to an approximate census tract FIPS code using the census tract corresponding to the geographical centroid of the zip code polygon.[24]

Finally, KN did not have either address or zip code information for the roughly 125,000 remaining RDD phone numbers (call this Subset 3). For these cases, the telephone exchange for each telephone number (i.e., the six digits making up the number's area code plus prefix) was used as the device for identifying an approximate census tract FIPS code. All of the census tracts overlapped by each active telephone exchange area—at the date of the recruitment attempt—were identified. (Directory-listed households in each identified census tract were enumerated separately.) The census tract with the largest number of directory-listed households was then designated as the "majority" census tract for that exchange. Each telephone number without address information was assigned to an approximate census tract FIPS code in this manner.

There are thus three sources of data for this particular analysis. Knowledge Networks first provided to us just their proprietary identity-protected street addresses (Subsets 1 and 2), with no other associated data, for geocoding. These addresses were associated with their census tract FIPS codes and returned to KN to have (a) the addresses removed, and (b) the sampling status and attrition history of each contact appended. Proxy case identifiers were generated and the files were returned to us for subsequent analysis. For initial RDD contacts without address information (Subset 3), KN facilitated the task of matching each RDD telephone exchange with the census tract that best approximates the bulk of the telephone numbers in that exchange, delivering proxy identifiers and census tracts FIPS codes, along with sampling status and attrition history for each of these cases.[25] Subsets 1, 2 and 3, with their corresponding status and attrition histories, were then combined into one huge file. Each record contains an 11-character census tract FIPS code and a set of five indicator variables that identify whether each initial contact survived through five attrition processes:

a.) initially recruited to the Panel
b.) initial profile data collected
c.) still a part of the active Panel when a sample was drawn for the particular study in question
d.) part of the sample drawn for our particular study
e.) responded to the invitation to participate (in a sufficiently complete fashion to be included in the final estimating sample).

In other work, we have examined the conditional retention propensities at each stage of attrition. For this illustrative analysis, however, we consider just the comprehensive selection between the households targeted in the original RDD recruitment attempts (524,890 telephone numbers) and the individual KN panel members whose responses are used in our analysis in the main paper corresponding to 1,801 actual respondents.

4.5.1.1 Comprehensive versus "end-stage" sample selection

It should be noted that in many other studies, sample selection bias is assessed only between stages (d.) and (e.) above. Researchers assume that the targeted households are a representative

[24] These links were accomplished using utilities provided within ArcView.

[25] Dale Kulp of Marketing Systems Group (MSG) generously provided the exchange/census-tract matching for this subsample. We note also that a very limited set of initial RDD contacts were lost from KN's archival records. However, we are confident that this block of lost data occurred essentially randomly. We have no recourse but to assume this loss was independent of any of the other general processes modeled here and to proceed without those data.

sample for the population, and response/nonresponse modeling explains survey recipients' decisions whether to participate in the survey. For this last transition, our response rate for those standing KN panelists actually invited to participate was a respectable 79 percent.

There has been an impetus over the last several years to acknowledge that selection should be addressed between the stage of the panel recruitment process that is truly a random sample from the population (or the closest thing to it—random digit dialed telephone contacts), and the final estimating sample, rather than just the "end-stage" sample selection between invited panelists and those who provide a set of responses used in the analysis. Our selection modeling responds to this growing requirement by modeling the selection of our 1,801 respondents used to produce our willingness to pay estimates, from the original set of almost 525,000 RDD contacts.

4.5.2 Census tract factors

We use the census tract FIPS codes for each tract to merge our data with the census tract factors resulting from the factor analysis described in Cameron and Crawford (2003). These factors capture variations in both short- and long-form census variables across tracts. These data consist of a set of 15 mutually orthogonal factors that capture approximately 88 percent of the variation—in a set of 95 variables—across the roughly 65,000 census tracts in the 2000 Census. Again using census tract identifiers (11-character FIPS codes), we then merge the fifteen factor scores with the original 524,890 RDD residential contact attempts.[26]

The use of local averages or aggregates in lieu of household-specific data is always a compromise. However, we argue that models based on at least some information about possible systematic differences across RDD contacts in the original contact group are preferable to the alternative of ignoring the endogenous selection process altogether.

There is a clear reason for preferring census tract factor scores to the alternative of using a vastly larger number of raw census variables. Many census variables are highly collinear, making it extremely difficult to tease out the distinct incremental effect of a difference in any one variable upon the outcome of interest (e.g., sample membership/non-membership). Estimated factors produced by factor analysis have the attractive property of being orthogonal by design. The factor scores span the same space as the much larger number of correlated variables upon which they are based, but they are uncorrelated, so their distinct effects can be identified more easily (if such effects are indeed present). It is our goal merely to control for systematic variation in attrition propensities, rather than to quantify the specific causes of attrition. Thus factor scores can be particularly valuable in selection correction models.

However, the downside of using estimated factor scores as explanatory variables is that they must typically be considered to be "estimated" quantities. Ordinarily, we are very concerned about this, since estimated quantities come with varying levels of precision. If we fail to recognize the estimated nature of factor scores, we will be understating the amount of noise in the overall model and distorting any hypothesis testing in any second-stage model which uses them. This is called the "estimated regressors" or "constructed regressor" problem in the econometrics literature. In this case, however, there is some basis for arguing that the estimated regressors problem is minimized. We are not using factor scores estimated for just the *sample* of

[26] For only a tiny minority of census tracts (i.e., less than 0.4%), it was not possible to construct a set of census tract factors. Thus we include an indicator variable, *census factors available*, that takes on a value of 1 if the census tract factors are available, and is zero otherwise.

census tracts represented in the RDD sample provided by Knowledge Networks. The factor scores used in this study are instead calculated *for the complete set of all census tracts in the U.S.* As such, our tract-level factor scores are technically not just *estimates* of the corresponding "population" factor score values, but are the calculated population values themselves (although only for the 2000 Census).

While our census tract factor scores therefore represent the "population values" of the tract-level factors, they are not the attributes of the specific individual who was contacted in the RDD sample. The census tract factors will be a better estimate of the individual's characteristics, the more homogeneous the population of the census tract. However, we are not able to control for the amount of noise introduced by using census tract characteristics as proxies for the individual characteristics that we would prefer to use if they were available.[27]

Across the universe of census tracts in the entire U.S., the mean and variance of the census tract factors should be zero and one, respectively, since the factor scores are standardized by the algorithm that calculates them. Modest departures from these standardized means and variances, for our half-million cases, reflect the slightly disproportionate presence of RDD contacts in some physical census tracts and also the approximations necessary to match telephone exchanges with the right census tracts.

4.5.3 Voting patterns in 2000 Presidential election

In many survey applications, especially if the research is intended to inform policy-making, we are concerned not only about whether sociodemographic groups are proportionately represented, but also about whether political constituencies are proportionately represented. To allow this question to be addressed in at least a rudimentary fashion, we have also merged into our data set, by county FIPS code, all of the available information at the county level about percentages of voters who voted for the Democratic candidate (Al Gore) and for the Green Party Candidate (Ralph Nader) in the 2000 Presidential election.[28]

4.5.4 County death rates

The salience of a survey about programs to reduce risks to life and health can be expected to be greater in communities where per-capita death rates are higher. We applied for and received authorization to work with the Compressed Mortality Files (CMF) from the National Center for Health Statistics (NCHS). We desire a proxy for county-level perceptions of death rates from each of the health threats featured in our survey (as well as some additional specific health threats featured in other surveys we conducted with Knowledge Networks during the same time period). The International Classification of Disease (ICD) codes were used to aggregate all counts of deaths from each health threat in each county in each year from 1988 to 2003. These

[27] Sworn employees of the U.S. Census Bureau can gain access to much more of the individual household data underlying the census tract totals. These data would allow the researcher to estimate the variance-covariance matrix for census variables within each census tract and would allow more rigorous corrections for this type of measurement error. This strategy, however, is still prohibitively difficult with current technologies and we do not have sworn Census employee status.

[28] Proportions in the "omitted category" voting for candidates other than the Republican candidate (George W. Bush) are assumed to be sufficiently small that little generality is lost by neglecting them. Presidential voting data are available in spreadsheet format from Leip (2003). However, the breakdown in votes for Alaska counties is not available. We thus include an indicator for election data availability, *vote percentage available*.

counts were summed across this range of years and then expressed as a fraction of the population in that county as of the 2000 Census. Thus, the variables are not annual average death rates, but they are roughly proportional to these rates, and arguably a reasonable proxy for the "recent prevalence of local death by each cause" for each of the 524,890 contacted households in Knowledge Networks' RDD pool.

4.5.5 County hospital densities

People may view preventative health programs such as those described in our choice sets as substitutes for treatment later on, should they contract these diseases. It is possible that potential respondents may find our survey more salient if they are worried about access to treatment, should they become sick in the future. The major purveyor of geographic information systems (GIS) data is ESRI. This company provides geocoded information about the locations of all hospitals in the U.S. We dropped closed facilities from the inventory and calculated the number of hospitals per unit area in each county, using ESRI's data for county areas. The ex ante expected effect of this "hospital density" variable is ambiguous, however. Greater access to hospitals may diminish people's interest in preventative care as a substitute for acute care or long-term care. However, a greater density of hospitals may reflect greater per-capita illness rates (demand for hospitals), which may increase the salience of a survey about health risks.

4.6 Sample selection assessment (comprehensive selection)

4.6.1 Binary probit selection model (n=524,890)

Table 4-3 displays parameter estimates from a binary probit model to explain membership in our estimating sample of the 1,801 respondents to our survey whose answers were sufficiently complete for analysis. Note that not every one of these respondents contributes a full set of five choices to the estimating data, however. Some respondents skip one or more of the choice tasks, other choices are disqualified on the basis of our minimal exclusion criteria (specifically, failure of the risk comprehension test, a preference for "Neither Program" when the sole explanation by the respondent is outright scenario rejection; or, that small fraction of choice sets where an unintended errors in the design of choice sets produced an illness profile depicting a modest extension to the individual's lifespan).

Nine of our fifteen census tract factors make statistically significant contributions to explaining the propensity of an RDD contact attempt to yield an eventual member of our estimating sample. Response propensities are systematically higher in census tracts where there are more "well-to-do-seniors," "elderly disabled," "rural faming, self-employed," "Native American" and "health care workers." Response propensities are systematically lower in census tracts where there are more "well-to-do prime-aged adults," "single renter twenty-somethings," "minority single moms," and "Asian, Hispanic, or language-isolated" households.

In contrast, controlling for sociodemographic factors, people's political ideologies at the county level, do not have a discernible effect on selection propensities. The coefficients on county voting percentages in the 2000 Presidential election, for Gore or for Nader (with Bush being the omitted share) are not statistically significant.

However, some of the different actual historical county death rates, conditional on data being available for an observation, do have systematic effects on selection propensities. Selection propensities are systematically higher, the higher the death rates from colon cancer, strokes,

Alzheimer's disease, and asthma. Selection propensities are systematically lower, however, the higher the historical county death rate from heart disease. Given the likelihood of multicollinearity in these various death rates, as a function of the age distribution already captured in this model by the sociodemographic factors embodied in the Census Tract factors, these estimated effects are incremental (i.e. in addition to what is captured by the age distribution information that has already been controlled for in the model).

Our goal is to produce the best possible fit in terms of point estimates of response probabilities or propensities, so we do not undertake to reduce the selection model to a parsimonious form that features only the most robustly statistically significant regressors.

4.6.2 Evaluating the potential for selection bias

The "second-stage" model in conventional selection correction models is typically an ordinary least squares regression. In such a case, it is standard to estimate the selection equation and the outcome equation simultaneously by maximum likelihood, and to focus on the estimated value of the correlation between the error terms in the two equations (as well as the differences in the parameter estimates for the outcome equation as a consequence of joint estimation).[29]

Our "outcome" model, however, is a fixed-effects conditional logit model for three-alternative choices. Logit models do not lend themselves readily to simultaneous estimation with selection models, and cross-equation error correlations have not yet been implemented in any econometric software packages. The main impediment is that logistic error distributions preclude error correlations between "choice equations." Simulation methods hold out some hope for researchers to overcome this limitation, but such models have not yet been implemented.

As a consequence of there being no readily available procedures that generalize conventional selection correction models to the context of ordinary or fixed-effects conditional logit models, we adopt a more *ad hoc* approach in our investigation of selection effects on the parameter estimates in our model of program choices. We propose instead a sensitivity assessment using the estimated selection probabilities from the model in Table 4-3.

We use fitted selection probabilities to investigate the possible effects of non-random sampling on the estimated parameters in our choice model.[30] We wish to know what *would have been* the vector of model parameters if each original RDD panel recruitment contact was equally likely (according to our selection equation) to show up in this particular estimating sample. Thus it is helpful to express all of the estimated probabilities as deviations from the "typical" selection probability in the RDD population. These normalized fitted selection probabilities are allowed to shift every outcome-model parameter. The baseline outcome-model parameter estimates then represents the "simulated" parameters for the counterfactual case where every respondent's chance of being in our estimating sample is equal to the central tendency in the original RDD pool (meaning that all deviations-from-typical are zero). This allows our basic parameter

[29] Sample selection models have been researched extensively. The seminal paper is Heckman (1979), and subsequent surveys of basic and alternative models have been provided by Vella (1998) and Das et al. (2003). The role of sociodemographic characteristics on response propensities has been considered by Hausman and Wise (1979), Ridder (1990), Lillard and Panis (1998), Fitzgerald et al. (1998a), Fitzgerald et al. (1998b), and Nicoletti and Peracchi (2005), among others.

[30] As mentioned above, the fitted selection probabilities from the selection equation represent estimated regressors in any second stage equation. A more-rigorous assessment would need to attempt corrections for the concerns raised in Pagan (1984) and Pagan and Nicholls (1984).

estimates to be systematically larger or smaller for observations with higher propensities to appear in the estimating sample, relative to their frequencies in the initial RDD contact pool.

We employ differences from the *median* response probability (as a measure of central tendency in the general population) so that the estimated utility parameters correspond to the simulated case where all response probabilities are exactly equal to the median in the population. We employ the median because the distribution is skewed, with a number of large positive outliers.

Model 4' in Table 4-4 illustrates the consequences of allowing the parameters of the model to vary according to the fitted probability that each respondent appears in our estimating sample. Only the coefficient on the first-order discounted sick-years term differs significantly with the fitted probability that the respondent shows up in our estimating sample. The greater the probability of being in our sample, relative to the median probability, the lesser the disutility the individual appears to experience from an increase in discounted sick-years. While the shift is statistically significant, comparison of Model 4' and Model 4 reveals the relatively minor difference in the magnitude of this key sick-years coefficient across individuals with different response propensities. The sick-years coefficient differs by less than 10%, without and with normalization on the median survey participation probability out of the 525,000 original RDD recruiting contacts. There are negligible differences in the other estimated coefficients.

4.7 Caveats concerning selection corrections

We have conducted a careful inquiry into the possibility of systematic selection in a sample drawn from the Knowledge Networks panel—between the original random-digit-dialed recruiting contact and a respondent's eventual participation in a particular research study sample. The most innovative feature of this sample-selectivity assessment/correction exercise is that we reach all the way back to the initial RDD recruiting contacts made to build the panel, rather than considering just the "end-stage" selectivity for the small subset of panelists actually invited to participate in this particular survey. We consider many characteristics of these panelists (proxied by the sociodemographic characteristics of the census tract where they live, or the voting patterns in their county or other "neighborhood" characteristics).

4.7.1 Group averages in lieu of individual data

It is worth reiterating that the use of census tract or county averages as proxies for individual values presents the usual "errors-in-variables" problem in selection-type models where the researcher must rely on these averages in lieu of the specific characteristics of each individual. Errors in regressors are typically expected to produce "errors-in-variables attenuation." As a consequence, failure to find statistically significant effects in these types of models does not necessarily mean they would *not* materialize if analogous individual-specific regressors were available.

4.7.2 Multiple stages of attrition

Candidate panelists either survive or do not survive each stage of attrition leading up to their final decision whether to participate in our survey after they have been invited to do so. At each stage, more individual-specific information is known, with the richest information being available for the invited panelists. However, conformable information is not available for the

individual RDD contact attempts, so no single comprehensive selection model can be estimated with more variables than we currently use to fit the participation probabilities using our selection model. A multi-selected model might be attempted, but do we do not pursue the possibility here, since the "outcome" model is a fixed effects conditional logit. Further analysis must await advances in selectivity correction technologies.

4.8 Conclusions

4.8.1 Some selection; not strongly related to health risk preferences

For the Knowledge Networks sample examined here, we find several statistically significant determinants of membership in the estimating sample, starting from the pool of over one-half million original RDD contacts. However, while there might be a "smoking gun" in this case, there appears to have been very little injury produced in terms of the distortions that heterogeneous participation propensities produce in the parameters of interest in the outcome model. Given that our outcome model is a fixed effects logit model with three alternatives, conventional selection correction models are not appropriate. However, we consider a less-sophisticated strategy to determine whether there is any systematic variation in parameters according to the estimated participation propensity (or probability) that each potential panelist, as an initial RDD recruiting contact attempt, ends up in the final estimating sample. We identify just one statistically significant systematic effect, but the magnitude of this effect is relatively small.

4.8.2 Little selection on political ideology (attitudes toward regulation)

On a final note, some audiences have expressed concern that the widely used Knowledge Networks panel may have either a "liberal bias" or a "conservative bias," but in Cameron and DeShazo (2010a), we have explored additional selectivity correction models to measure respondents' answers, on a separate survey sample drawn from the same panel, to a question concerning the proper role of government in regulating environmental, health, and safety risks. Our model in that case does permit a standard maximum likelihood selection correction model (provided we treat the rating in each person's answer as a continuous variable). Our results do not support the presence of either a liberal or a conservative bias (to the extent that this would be correlated with attitudes toward the proper role of government in risk regulation) as a result of sample selection. In particular, controlling for sociodemographics, there is a somewhat lower overall response probability for panelists from counties where a higher proportion of votes in the 2000 Presidential election went to Gore. However, this effect is offset to a considerable extent by a slightly higher overall response probability for panelists from counties where a higher proportion of votes went to Nader.

Overall, our results can probably be characterized as reassuring news for researchers who have used (or who contemplate using) the Knowledge Networks panel for policy-oriented research. These finding are also welcome, presumably, for policy makers who need to rely on survey-based research to support their decisions and who would prefer to have demand data from a sample with no particular ideological biases. Of course, these results are suggestive, rather than conclusive, but outcome discerned in our analysis is better than the alternative.

4.9 Section 4 Tables

Table 4-1 Sample versus population characteristics
(See Section 4.3.1; with scenario adjustment/rejection controls, the sample can be expanded to 2,407 respondents.)

	Sample (n=1,801 Individuals)	2000 U.S. Census
Age	**%**	% of 25+ pop
25 to 34	16.7	22
35 to 44	22.8	25
45 to 54	21.5	21
55 to 64	17.7	7
65 to 74	14.3	6
75 and older	6.9	10
Income		% of hhlds
Less than $10,000	5.7	9.5
$10,000 to $15,000	6.1	6.3
$15,000 to $20,000	5.1	6.3
$20,000 to $25,000	6.4	6.6
$25,000 to $30,000	6.6	6.4
$30,000 to $40,000	16.2	6.4
$40,000 to $50,000	13.2	5.9
$50,000 to $60,000	11.3	10.7
$60,000 to $75,000	10.9	9.0
$75,000 to $100,000	10.2	10.4
$100,000 to $125,000	4.1	10.2
More than $125,000	4.3	5.2
Female	0.52	0.51

Table 4-2 Assessing the impact of sample exclusion criteria
(See Section 4.3.2. Estimating sample when these comparison were made was = Sample 4)

Sample: Parameter	(1) none	(2) by[a]	(3) by,cr	**(4) by,cr,wk**	(5) by,cr,wk,60	(6) by,cr,wk,80	(7) by,cr,wk,100
β_0	.01848	.01833	.01986	**.01392**	.01414	.01403	.01416
	(14.56)***	(14.10)***	(13.49)***	**(9.48)*** **	(9.47)***	(9.25)***	(9.06)***
α_{10}	29.07	27.58	14.24	**-46.78**	-51.33	-49.87	-50.8
	(3.87)***	(3.62)***	(1.70)*	**(5.40)*** **	(5.84)***	(5.58)***	(5.55)***
α_{20}	.6698	-.9993	-1.835	**-16.72**	-17.93	-19.31	-19.82
	(0.08)	(0.12)	(0.20)	**(1.79)* **	(1.89)*	(2.01)**	(2.00)**
α_{30}	-146.9	-143.9	-342.6	**-564.9**	-637.9	-647.6	-631.8
	(0.98)	(0.95)	(2.04)**	**(3.17)*** **	(3.53)***	(3.49)***	(3.31)***
α_{31}	9.348	9.368	16.15	**19.77**	21.91	21.59	20.7
	(1.52)	(1.51)	(2.36)**	**(2.73)*** **	(2.99)***	(2.89)***	(2.70)***
α_{32}	-.0826	-.08379	-.1469	**-.1814**	-.1968	-.1904	-.1837
	(1.40)	(1.40)	(2.24)**	**(2.62)*** **	(2.81)***	(2.67)***	(2.53)**
α_{40}	86.11	80.27	129.1	**196.1**	222.7	231.7	230.3
	(1.24)	(1.14)	(1.65)*	**(2.38)** **	(2.67)***	(2.71)***	(2.62)***
α_{41}	-4.503	-4.346	-6.05	**-7.575**	-8.396	-8.473	-8.325
	(1.58)	(1.51)	(1.90)*	**(2.26)** **	(2.47)**	(2.45)**	(2.35)**
α_{42}	.04108	.0402	.05665	**.07214**	.07818	.07742	.07667
	(1.48)	(1.43)	(1.83)*	**(2.22)** **	(2.38)**	(2.32)**	(2.25)**
α_{50}	-17.56	-13.88	16.17	**106.5**	100.9	98.27	71.73
	(0.29)	(0.23)	(0.23)	**(1.46)**	(1.37)	(1.30)	(0.92)
α_{51}	-1.417	-1.563	-2.466	**-4.588**	-4.367	-4.308	-3.422
	(0.59)	(0.65)	(0.91)	**(1.61)**	(1.52)	(1.47)	(1.13)
α_{52}	.02488	.02692	.03611	**.05684**	.05556	.05512	.04809
	(1.11)	(1.18)	(1.43)	**(2.13)** **	(2.06)**	(2.01)**	(1.71)*
Alternatives	35151	34155	27795	22,560	21855	21030	19881

Log L	-18267.16	-17687.02	-14446.62	**-11687.98**	-11310.84	-10884.97	-10288.77

[a] **Key to exclusion criteria:** "by" = (bad year) = choice did not involve an (erroneously designed) minor life extension from the illness experience; "cr" = (comprehend risk) = respondent passed simple risk comprehension question at end of risk tutorial; "wk" = (weak scenario rejection) = choice of Neither Program not explained solely by "I did not believe the programs would work" (i.e. outright scenario rejection); "60" = aggregate time on all five program choice tasks at least 60 seconds (e.g. average time at least 12 seconds per choice set); analogously for "80" and "100." The most substantial incremental impact is associated with the "wk" criterion. In our main paper, only the "by" exclusion criterion is now employed, because other observations can be salvaged via the use of scenario adjustment/rejection corrections.

Table 4-3 Probit model used to calculate fitted selection probabilities

(See Section 4.6.1. Sample size=524,890; analogout models have been employed when fewer exclusion criteria and scenario adjustment/rejection controls have been employed. This is an example of the type of model used in our earliest published papers using these data, where we used three main exclusion criteria.)

Description of variable	Coefficient
Orthogonal factors from factor analysis of 2000 Census tract data	
Census factors available?	-0.677
	(9.28)***
"well-to-do prime aged adults"	-0.0862
	(8.01)***
"well-to-do seniors"	0.0349
	(3.52)***
"single renter twenties"	-0.0323
	(3.62)***
"unemployed"	-0.0147
	(1.55)
"minority single moms"	-0.0206
	(1.96)**
"thirty-somethings"	-0.0128
	(1.17)
"working-age disabled"	0.0036
	(0.36)
"some coll, no graduation"	-0.0082
	(0.86)
"elderly disabled"	0.0229
	(2.61)***
"rural farming, self-employed"	0.0386
	(2.77)***
"low-mobility stable neighborhood"	-0.0003
	(0.03)
"Native American"	0.0235
	(2.09)**
"female"	0.0001
	(0.01)
"health-care workers"	0.0337
	(3.22)***
"asian-hisp, language-isolated"	-0.0530
	(4.45)***

Voting patterns in 2000 Presidential Election	
County vote percentages available?	-1.02
	(11.36)***
Percent voting for Gore (liberal bias?)	-0.105
	(1.02)
Percent voting for Nader (liberal bias?)	0.324
	(0.58)
Average county death rates, Compressed Mortality Files (1989-2003)	
County death rate available for lung cancer available?	0.163
	(1.39)
County death rate available for skin cancer available?	0.0956
	(0.27)
County death rate available for heart attacks available?	0.483
	(1.59)
County death rate available for asthma available?	-0.157
	(1.92)*
County death rate from breast cancer	8.99
	(0.06)
County death rate from prostate cancer	1790.
	(0.47)
County death rate from colon cancer	570.
	(2.15)**
County death rate from lung cancer	-463.
	(0.79)
County death rate from skin cancer	-95.0
	(0.28)
County death rate from heart disease	-36.6
	(3.06)***
County death rate from heart attacks	-21.0
	(0.41)
County death rate from strokes	8.99
	(2.57)**
County death rate from respiratory disease	12.1
	(1.09)
County death rate from traffic accidents	-23.9
	(0.32)
County death rate from diabetes	-59.5
	(1.02)
County death rate from Alzheimer's disease	161.
	(2.77)***
County death rate from leukemia	-263.

	(1.46)
County death rate from leukemia in children	92.5
	(0.07)
County death rate from asthma	391.
	(2.08)**
County death rate from asthma in children	-948.
	(0.38)
County death rate from cancers in general	-36.8
	(0.99)
Hospital density calculated from ESRI GIS shapefiles for US hospitals	
County hospital density available?	-0.0097
	(0.34)
Hospital density	-0.0001
	(0.48)
Constant	-1.49
	(3.05)***
Number of observations (Knowledge Networks initial recruiting contact attempts by random digit dialing)	524,890
Number of respondents in estimating sample	1,801

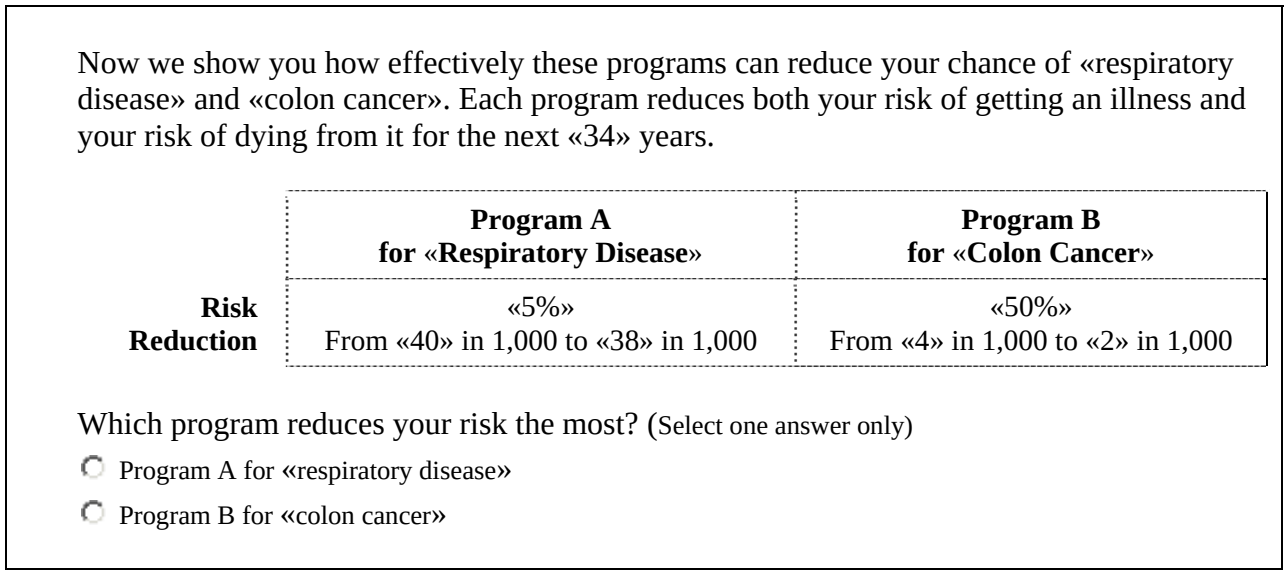
Absolute value of z statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4-4 Sensitivity of utility parameters to response probability
(See Section 4.6.2. Illustration using preliminary fixed effects logit models; Individuals = 1,801; choice sets = 7,520; with exclusion restrictions in lieu of extensive scenario adjustment/rejection controls.)

	Model 4	Model 4'
(Parameter) Variable	Simple logs	w/ P(select)
$(\beta_0)\left(\left(Y_i - c_i^j\right)^{(.42)} cterm_i^j - \left(Y_i\right)^{(.42)} yterm_i^j\right)$	.01394 (10.47)***	.01389 (10.43)***
$(\alpha_{10})\Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$	-26.44 (4.61)***	-28.78 (4.95)***
$\ldots(\alpha_{13})\left[P(sel_i) - \overline{P}\right]\Delta\Pi_i^{jS}\left[\log\left(pdvi_i^j + 1\right)\right]$	-	3.282[a] (2.28)**
$(\alpha_{20})\Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	-22.46 (2.41)**	-22.53 (2.42)**
$(\alpha_{30})\Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-27.73 (5.60)***	-27.64 (5.58)***
Log L	-11719.76	-11717.04

[a] Only this interaction term bears a coefficient that is statistically significantly different from zero. Similar results are obtained for the model with higher-order terms and age effects.

4.10 Section 4 Figures



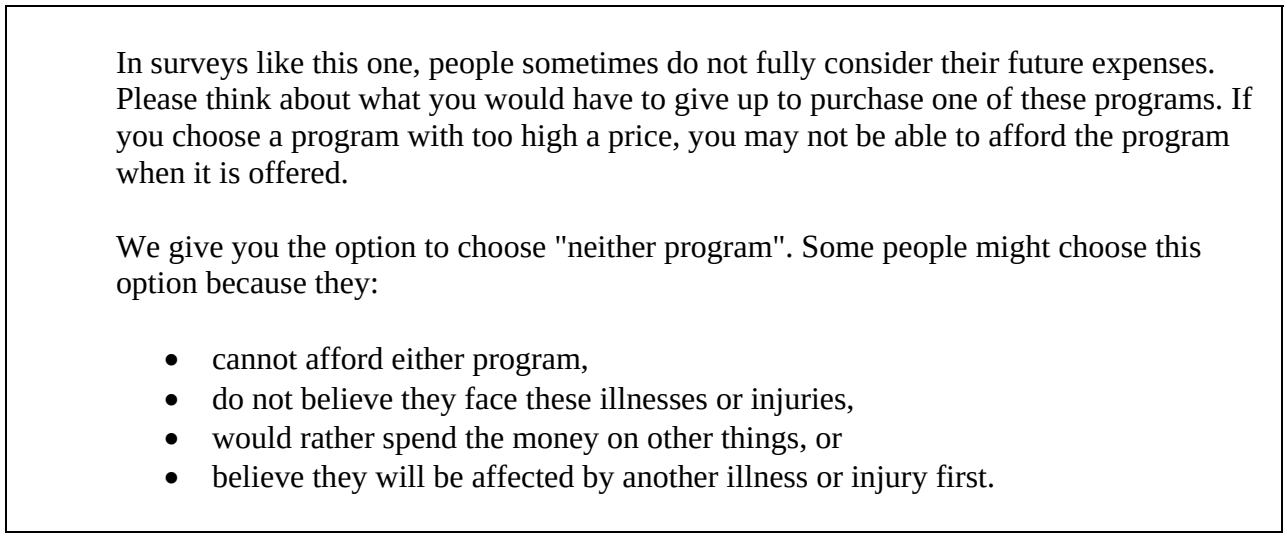
Now we show you how effectively these programs can reduce your chance of «respiratory disease» and «colon cancer». Each program reduces both your risk of getting an illness and your risk of dying from it for the next «34» years.

	Program A for «Respiratory Disease»	**Program B for «Colon Cancer»**
Risk Reduction	«5%» From «40» in 1,000 to «38» in 1,000	«50%» From «4» in 1,000 to «2» in 1,000

Which program reduces your risk the most? (Select one answer only)

- ○ Program A for «respiratory disease»
- ○ Program B for «colon cancer»

Figure 4-1 Risk reduction information in choice scenarios



In surveys like this one, people sometimes do not fully consider their future expenses. Please think about what you would have to give up to purchase one of these programs. If you choose a program with too high a price, you may not be able to afford the program when it is offered.

We give you the option to choose "neither program". Some people might choose this option because they:

- cannot afford either program,
- do not believe they face these illnesses or injuries,
- would rather spend the money on other things, or
- believe they will be affected by another illness or injury first.

Figure 4-2 Cheap talk wording

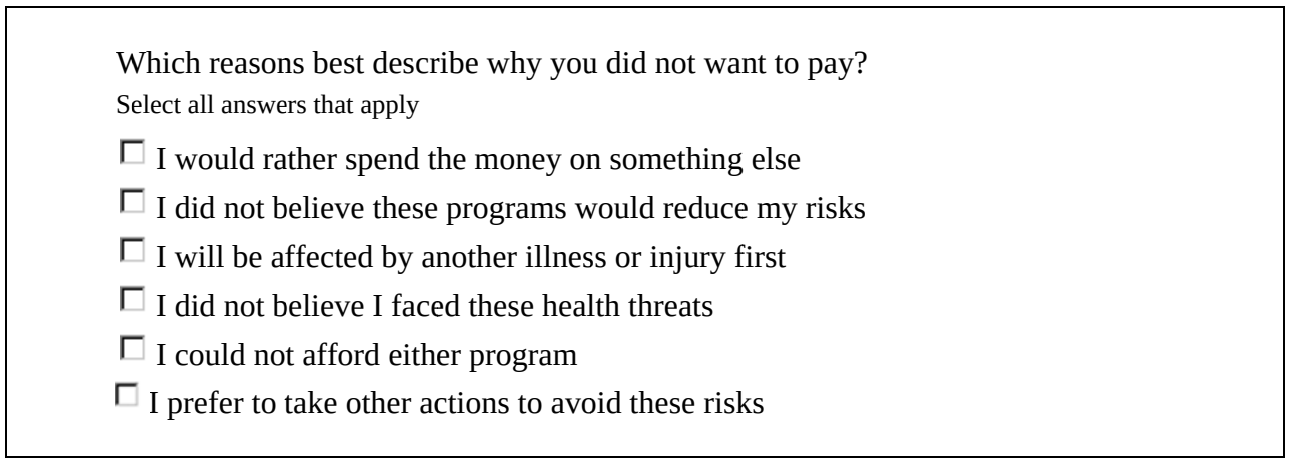


Figure 4-3 Reasons for choosing neither program

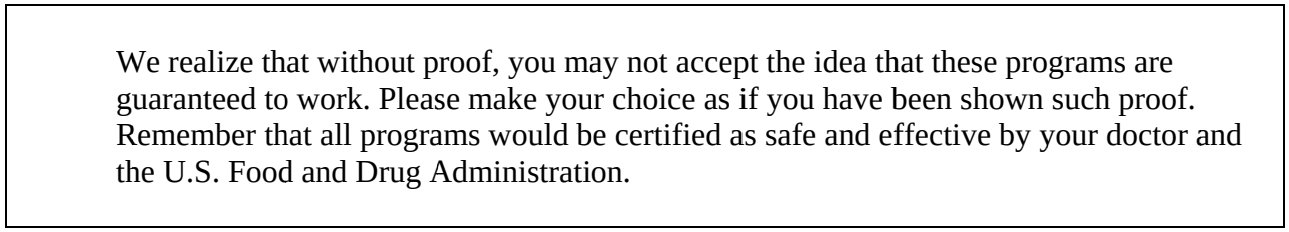


Figure 4-4 Assurances of efficacy

5 Model, Estimation and Alternative Analyses

Equation Section (Next)

In this section, we provide detailed descriptions of how we arrive at the estimating specifications used in our model. For readers who may be unfamiliar with econometric methods for discrete choice data, we also provide an outline of how Stata's fixed effects conditional logit model is specified and estimated.

The illness profiles presented to respondents in each of the choice scenarios are variations on the forms shown in Figure 5.1. This figure shows three illustrations of illness profiles like the ones we describe to respondents in the choice scenarios with which they are presented. Each scenario involves only a single spell of illness, rather than multiple spells, which were beyond the scope of the study. The individual's current age and gender define the number of remaining nominal life-years at stake. The individual is informed that they face an existing baseline risk of each featured illness profile. The programs they are asked to consider do not change the time profile of the illness in question, only its likelihood of occurring.

The interventions in our choice scenarios merely *alter the probability* of experiencing a given illness profile. Choice scenarios where the interventions specifically *alter the time profiles* of the illnesses described are left for future research. Nevertheless, our fitted models can permit an approximate simulation of a change in the time profile of an illness in two stages—first, via a reduction in the probability of the original illness profile to zero, and second, a corresponding increase in the probability of the new, different illness profile.

5.1 Derivation of the estimating forms of the model

Let the superscript A denote "under Program A" while N denotes "with no program". Let the superscript H denote "if the respondent remains healthy," while S denotes "if the respondent gets sick from this health threat." We suppress the i subscripts for individuals and write indirect utility levels as a function of net income and health status in each future period (already denoted *relative* to their current health) as follows.

$$\begin{aligned} V_t^{AH} &= f\left(netY_t\right) + \varepsilon_t^{AH} \\ V_t^{AS} &= f\left(netY_t\right) + \alpha_1 1\left(illness_t\right) + \alpha_2 1\left(recovered_t\right) + \alpha_3 1\left(lost\ life\text{-}year_t\right) + \varepsilon_t^{AS} \\ V_t^{NH} &= f\left(netY_t\right) + \varepsilon_t^{NH} \\ V_t^{NS} &= f\left(netY_t\right) + \alpha_1 1\left(illness_t\right) + \alpha_2 1\left(recovered_t\right) + \alpha_3 1\left(lost\ life\text{-}year_t\right) + \varepsilon_t^{NS} \end{aligned} \tag{0.1}$$

For future period t, we can write the difference in expected utility with program A and with no program (N). Note that across the papers in this project, we sometimes use lower-case π to denote the probability of getting sick, whereas this exposition uses the upper-case version of the notation:

$$\left[\left(1-\Pi^{AS}\right)V_t^{AH} + \left(\Pi^{AS}\right)V_t^{AS}\right] - \left[\left(1-\Pi^{NS}\right)V_t^{NH} + \left(\Pi^{NS}\right)V_t^{NS}\right] \tag{0.2}$$

To explain a decision taken today, based on the stream of future *differences* in expected indirect utilities across the two alternatives, these future quantities must be discounted back to the present.

The fact that net income and health status are assumed to be approximately level within each of the four different health states permits us to reverse the order of discounting and the taking of expectations. We can work in terms of the present discounted *time* in each health state, and simply multiply this by the utility of net income in that interval and by the (dis)utility of health status in that interval. We assume simple exponential discounting where the discount factor is $\delta^t = (1+r)^{-t}$. Each summation in the following terms runs from the present to the end of the individual's nominal lifespan.

$$\begin{aligned} pdve &= \sum \delta^t 1(\, pre\text{-}illness_t) \\ pdvi &= \sum \delta^t 1(illness_t) \\ pdvr &= \sum \delta^t 1(recovered_t) \\ pdvl &= \sum \delta^t 1(lost\ \ life\text{-}year_t) \end{aligned} \tag{0.3}$$

For convenience, we define two other types of present discounted time intervals, $pdvp = pdve + pdvr$, which captures just the time where the individual is neither sick nor dead, and $pdvc = pdve + pdvi + pdvr + pdvl$, which corresponds to the entire remainder of the individual's nominal lifespan.

We will now develop, separately, the "present discounted expected" form of the three parts of the indirect utility function: the net income terms, the health status terms, and the error term. Fortunately, we find no strong evidence that the marginal utility of net income depends on these probabilistic future health states (or vice versa) in any of the models explored in the main paper. In other work, we find some evidence of the dependence of the marginal utility of net income on current health, but this is the numeraire health state in the main paper.

5.1.1 Development of the net income term

Table 5-1 shows the pattern of income and program costs over the individual's future life-years, as a function of whether the program is selected and whether he/she gets sick. The net income level, $netY_t$, will differ according to the type of health state, whether the program is currently being paid for, and whether the individual gets sick or stays healthy:

We can make use of our notation for discounted future time intervals, plus the pattern of net income amounts under the four different outcomes as displayed in Table 5-1, to specify the discounted future expected utility from net income (noting that $pdve + pdvi + pdvr = pdvc - pdvl$). The parameters γ_1 and γ_2 allow different assumptions about the fraction of the respondent's current income that would be received if they are sick or dead from the illness in question. The parameters γ_3 and γ_4 allow varying assumptions about what fraction of risk reduction costs the respondent would be obliged to pay when sick or dead. Note that in any model wherein indirect utility is not a linear function of net income, it appears to be necessary (to make it straightforward to solve for c) to limit the coefficients γ_3 (the fraction of program costs

paid while sick) and γ_4 (the fraction of program costs paid after death) to take on only the values 0 or 1. Otherwise, it is prohibitively difficult to solve the utility-difference function for an expression for willingness to pay (WTP).

The discounted future expected utility from net income and future health states can then be written as follows:

$$\begin{bmatrix} \left(1-\Pi^{AS}\right) f\left(Y-c\right) pdvc \\ +\left(\Pi^{AS}\right)\left[f\left(Y-c\right) pdve + f\left(\gamma_1 Y-\gamma_3 c\right) pdvi + f\left(Y-c\right) pdvr + f\left(\gamma_2 Y-\gamma_4 c\right) pdvl\right] \end{bmatrix} \\ -\begin{bmatrix} \left(1-\Pi^{NS}\right) f\left(Y\right) pdvc \\ +\left(\Pi^{NS}\right)\left[f\left(Y\right) pdve + f\left(\gamma_1 Y\right) pdvi + f\left(Y\right) pdvr + f\left(\gamma_2 Y\right) pdvl\right] \end{bmatrix} \tag{0.4}$$

Distribute the probabilities and rearrange to get:

$$\begin{aligned} & f\left(Y-c\right) pdvc - \Pi^{AS} f\left(Y-c\right) pdvc + \Pi^{AS} f\left(Y-c\right)\left(pdve + pdvr\right) \\ & \quad + \Pi^{AS} f\left(\gamma_1 Y-\gamma_3 c\right) pdvi + \Pi^{AS} f\left(\gamma_2 Y-\gamma_4 c\right) pdvl \\ & - f\left(Y\right) pdvc + \Pi^{NS} f\left(Y\right) pdvc - \Pi^{NS} f\left(Y\right)\left(pdve + pdvr\right) \\ & \quad - \Pi^{NS} f\left(\gamma_1 Y\right) pdvi - \Pi^{NS} f\left(\gamma_2 Y\right) pdvl \end{aligned} \tag{0.5}$$

We have noted that each of γ_3 and γ_4 may take on only the values of 0 or 1. If $\gamma_3 = 1$, then γ_1 must also be 1, so that $\gamma_1 Y - \gamma_3 c = Y - c$ and this term can be grouped with the other terms in $Y - c$. Likewise, if $\gamma_4 = 1$, then we must have $\gamma_2 = 1$ so that this term may also be included in the same group of terms. However, if $\gamma_3 = 0$, then $0 \le \gamma_1 \le 1$ can be accommodated and the term $\gamma_1 Y - \gamma_3 c = \gamma_1 Y$ can be grouped with the other term in $\gamma_1 Y$.

Gathering the terms in $f\left(Y-c\right)$, $f\left(Y\right)$, $f\left(\gamma_1 Y\right)$ and $f\left(\gamma_2 Y\right)$ and simplifying allows equation (0.5) to be written as follows. (Note that the fact that γ_3 and γ_4 can take on only the values of zero or one means that they can be used as indictors to switch on and off the presence of terms in $pdvi$ and $pdvl$.)

$$\begin{aligned} & f\left(Y-c\right)\left[\left(1-\Pi^{AS}\right) pdvc + \Pi^{AS}\left(pdve + \gamma_3 pdvi + pdvr + \gamma_4 pdvl\right)\right] \\ & \quad + f\left(Y\right)\left[(-1)\left\{\left(1-\Pi^{NS}\right) pdvc + \Pi^{NS}\left(pdve + pdvr\right)\right\}\right] \\ & + f\left(\gamma_1 Y\right)\left[\left((1-\gamma_3)\Pi^{AS} - \Pi^{NS}\right) pdvi\right] \\ & + f\left(\gamma_2 Y\right)\left[\left((1-\gamma_4)\Pi^{AS} - \Pi^{NS}\right) pdvl\right] \end{aligned} \tag{0.6}$$

To permit the use of further abbreviations for the terms which multiply the function $f\left(\cdot\right)$ in each of its four forms, we denote the four terms in square brackets in equation (0.6) as:

$$
\begin{aligned}
cterm &= \left(1-\Pi^{AS}\right) pdvc + \Pi^{AS}\left(pdve + \gamma_3 pdvi + pdvr + \gamma_4 pdvl\right) \\
yterm1 &= (-1)\left\{\left(1-\Pi^{NS}\right) pdvc + \Pi^{NS}\left(pdve + pdvr\right)\right\} \\
yterm2 &= \left((1-\gamma_3)\Pi^{AS} - \Pi^{NS}\right) pdvi \\
yterm3 &= \left((1-\gamma_4)\Pi^{AS} - \Pi^{NS}\right) pdvl
\end{aligned}
\tag{0.7}
$$

In the definitions in (0.7), it should be clear that depending upon whether γ_3 and γ_4 are either 0 or 1, two terms in $cterm$ and one term each in $yterm2$ and $yterm3$ will be switched either on or off, accordingly.

For estimation of the parameters of the model, we use these components to construct the net-income-related variable in the formula for the discounted expected utility difference:

$$
bXterm = f\left(Y-c\right) cterm + f\left(Y\right) yterm1 + f\left(\gamma_1 Y\right) yterm2 + f\left(\gamma_2 Y\right) yterm3 \tag{0.8}
$$

where $bXterm$ uses the indicator X to signify models with different functions $f(\cdot)$. The estimated coefficient on this variable can be interpreted as the marginal indirect utility associated with transformed net income, $f(Y)$, which has been factored out of each term involving $f(\cdot)$ on the right-hand-side of equation (0.8).

5.1.2 Development of the health-state-related term

Table 5-2 lays out the pattern of *utility* levels as a function of health states over the individual's remaining life-years, according to whether he/she suffers the illness profile in question. We assume that our subjects view future health states, when "healthy" or "sick," as being unaffected by whether Program A or No Program is selected (given that there is merely a *lesser* chance of getting sick if the risk reduction program is selected, not a *zero* chance). All that is affected by Program A is the *risk* of suffering this illness profile, not the illness profile itself. Unlike the net income profiles, therefore, the "net health" profile over time depends only on whether the individual gets sick.

Written in its extensive form the difference in discounted expected health states between Program A and no program is given by:

$$
\begin{aligned}
&\left[\begin{array}{l}\left(1-\Pi^{AS}\right)\left[\alpha_0 pdve + \alpha_0 pdvi + \alpha_0 pdvr + \alpha_0 pdvl\right] \\ +\left(\Pi^{AS}\right)\left[\alpha_0 pdve + \alpha_1 pdvi + \alpha_2 pdvr + \alpha_3 pdvl\right]\end{array}\right] \\
-&\left[\begin{array}{l}\left(1-\Pi^{NS}\right)\left[\alpha_0 pdve + \alpha_0 pdvi + \alpha_0 pdvr + \alpha_0 pdvl\right] \\ +\left(\Pi^{NS}\right)\left[\alpha_0 pdve + \alpha_1 pdvi + \alpha_2 pdvr + \alpha_3 pdvl\right]\end{array}\right]
\end{aligned}
\tag{0.9}
$$

Distributing the probability terms and simplifying yields:

$$\left(\Pi^{AS}-\Pi^{NS}\right)\begin{bmatrix}\left(\alpha_0-\alpha_0\right)pdve\\+\left(\alpha_1-\alpha_0\right)pdvi\\+\left(\alpha_2-\alpha_0\right)pdvr\\+\left(\alpha_3-\alpha_0\right)pdvl\end{bmatrix} \tag{0.10}$$

If we normalize future health-related utility on the individual's *status quo* health state, equivalent to setting $\alpha_0 = 0$, and express the change in the risk of the illness profile due to Program A as $\Delta\Pi^{AS} = \Pi^{AS} - \Pi^{NS}$, we can write this term more simply as:

$$\left(\alpha_1 pdvi + \alpha_2 pdvr + \alpha_3 pdvl\right)\Delta\Pi^{AS} = \alpha \text{term}\,\Delta\Pi^{AS} \tag{0.11}$$

Here, the estimated α_j parameters are the (dis)utilities from one unit of time in each adverse health state, *relative to* the individual's current pre-illness health status. This normalization is particularly convenient. However, it imposes some strong assumptions which we explore in other work, where we allow these marginal disutilities of adverse future health states to depend upon current morbidities and comorbidities, and upon subjective risks for the health problem in question and other major types of health risks. The marginal disutilities estimated in our basic models must be interpreted as averages, across the current population distribution of health states and health outlooks, for the U.S. population 25 years and older, across the range of health threats names in our study.

5.1.3 Development of the error term

For completeness, the assumed independent and identically error terms in each of the four variants of indirect utility in each future period are combined in a similar fashion:

$$\left[\left(1-\Pi^{AS}\right)\varepsilon_t^{AH} + \left(\Pi^{AS}\right)\varepsilon_t^{AS}\right] - \left[\left(1-\Pi^{NS}\right)\varepsilon_t^{NH} + \left(\Pi^{NS}\right)\varepsilon_t^{NS}\right] \tag{0.12}$$

When discounted back to the present, we assume the resulting differences in expected error terms (across the healthy and sick outcomes) are cooperative in being distributed in a manner consistent with the assumptions necessary for the use of McFadden's conditional logit choice model.

5.1.4 The difference in discounted expected utilities that drives choices

We can now assemble the discounted net income terms, the discounted health state terms, and the discounted error terms to yield the difference in discounted expected utilities that is assumed to drive the individual's choice between Program A and "No program."

$$\Delta PDV\left(E\left[V\right]\right)=\left\{\begin{array}{l} f\left(Y-c\right)cterm \\ \quad +f\left(Y\right)yterm1 \\ +f\left(\gamma_1 Y\right)yterm2 \\ +f\left(\gamma_2 Y\right)yterm3 \end{array}\right\}+\alpha term\,\Delta\Pi^{AS}+\varepsilon \qquad (0.13)$$

where $aterm=\left[\alpha_1 pdvi+\alpha_2 pdvr+\alpha_3 pdvl\right]$, to simplify the notation in what follows.[31] This is the basis for the estimating equations used in our papers.

Generalization to the case of three alternatives simply means we introduce a second "difference" equation analogous to equation (0.13), but for risk reduction Program B, relative to "No Program." Program costs and the size of the risk reduction, as well as the relevant illness profile, will differ between the two programs. For the "Neither program" alternative, of course, the "difference relative to Neither Program" is zero for all variables. There is no difference in net income because program costs are not incurred, and the term involving the health profile is zero because there is no reduction in the risk of experiencing that profile (i.e. $\Delta\Pi^{jS}=0$). The health risk is still present, but since neither program is selected, no reduction in risk is achieved.

All that remains is to choose a specific functional form for $f\left(\cdot\right)$ and to decide whether preferences are homogeneous or whether the data suggest that they should be specified as heterogeneous (i.e. a function of observable individual attributes). In our main paper, we eventually depart from this model based on future individual per-period health state utilities. Instead, we allow individuals' decisions to be based directly on "present discounted time in future adverse health states" as the proximal determinants of choice. We consider nonlinear forms in $pdvi$, $pdvr$, and $pdvl$, and find that a flexible translog-type functional form seems to provide the best fit to the choice data among familiar and easily estimated forms.

The data also suggest that the function $f\left(\cdot\right)$ should be nonlinear. We have explored quadratic forms, square root forms, and Box-Cox-type forms with a transformation parameter of 0.45, determined via a line-search. The quadratic form is the most general, but it involves one more parameter and it also permits marginal utility to go negative at extreme values of net income in some models with heterogeneous marginal utilities. The square root form is very close to the Box-Cox transformation with a parameter of 0.5, but reviewers of our early results have suggested that the 0.45 parameter may be preferable. In our main paper, we treat this parameter as a known constant, rather than estimating it using a fixed effects conditional logit model with a nonlinear-in-parameters "index" ($x\beta$ term) since such a model is not readily available. Treating this value of the parameter as fixed is certainly no worse than using a linear or logarithmic specification and implicitly assuming a Box-Cox transformation parameter that is fixed at one or zero.

The systematic portion of equation (0.13), provided it can be written as a linear-in-*parameters* function of variables constructed from our data, can be interpreted as the $x\beta$ term in the standard conditional logit (and fixed effects conditional logit) models that we use to estimate the parameters of our models. In other work, we are developing models which permit

[31] We generalize our specification so that utility is not merely linear in the level of discounted future health-state years, so $\alpha term$ will be more complex than this. It will involve nonlinear and interaction terms, as well as heterogeneity in some of the parameters with respect to respondent age.

nonlinearities in parameters in the logit index, in particular to accommodate estimated values of the discounting parameter. We treat the discount parameter as fixed in the main paper, although we are careful to explore the consequences of alternative assumptions about its magnitude.

5.1.5 Solving for *WTP*

Once the indirect utility parameters have been estimated using respondents' choices, the next step is to solve for the value of c that makes the individual just indifferent between paying for the program and getting the benefits, or not paying for the program and doing without the benefits. Recall that c is the annual payment for the risk-reduction program, assumed to be paid only if the individual is not currently afflicted by the illness in question or prematurely dead from the same illness.

In general, equation (0.13) can be solved for this maximum annual willingness to pay (while not sick or dead). For ease of verification, we show each step below:

$$
\begin{aligned}
0 &= f(Y-c^*)\,cterm + f(Y)\,yterm1 + f(\gamma_1 Y)\,yterm2 + f(\gamma_2 Y)\,yterm3 + \alpha term\,\Delta\Pi^{AS} + \varepsilon \\
-f(Y-c^*)\,cterm &= f(Y)\,yterm1 + f(\gamma_1 Y)\,yterm2 + f(\gamma_2 Y)\,yterm3 + \alpha term\,\Delta\Pi^{AS} + \varepsilon \\
f(Y-c^*) &= \frac{-1}{cterm}\left(f(Y)\,yterm1 + f(\gamma_1 Y)\,yterm2 + f(\gamma_2 Y)\,yterm3 + \alpha term\,\Delta\Pi^{AS} + \varepsilon \right) \quad (0.14) \\
Y - c^* &= f^{-1}\left\{ \frac{-1}{cterm}\left(f(Y)\,yterm1 + f(\gamma_1 Y)\,yterm2 + f(\gamma_2 Y)\,yterm3 + \alpha term\,\Delta\Pi^{AS} + \varepsilon \right) \right\} \\
c^* &= Y - f^{-1}\left\{ \frac{-1}{cterm}\left(f(Y)\,yterm1 + f(\gamma_1 Y)\,yterm2 + f(\gamma_2 Y)\,yterm3 + \alpha term\,\Delta\Pi^{AS} + \varepsilon \right) \right\}
\end{aligned}
$$

We set the symmetric error term ε to its expected value of zero and ignore the variance of this error term in our calculations (although a logistic error with unit variance could readily be incorporated into our simulations). Our simulations focus on the variance-covariance matrix for the fitted parameter vector from our estimated models.

5.1.5.1 Special case: the linear form

It can be helpful to consider the *WTP* formula if the income function is $f(Y) = \beta Y$, because this simpler form aids in developing a clear intuition about the determinants of *WTP*. For the i^{th} individual and the illness profile for that individual that would be addressed by Program A (which would reduce the risk of getting sick with that illness by $\Delta\Pi_i^{AS}$), the maximum annual *WTP* during non-sick and non-dead years will be:

$$
c_i^{A*} = Y_i - \left(\frac{-1}{\beta cterm_i^A} \right) \left(\begin{array}{l} \beta Y_i\, yterm1_i^A + \beta(\gamma_1 Y)\, yterm2_i^A + \beta(\gamma_2 Y)\, yterm3_i^A \\ +\alpha_1 \left\{ \Delta\Pi_i^{AS}\, pdvi_i^A \right\} + \alpha_2 \left\{ \Delta\Pi_i^{AS}\, pdvr_i^A \right\} + \alpha_3 \left\{ \Delta\Pi_i^{AS}\, pdvl_i^A \right\} + \varepsilon_i^A \end{array} \right) \quad (0.15)
$$

The individual expects to pay this constant annual cost of the program in periods when he or she is initially healthy or in a recovered/remission state, but the parameters $\gamma_3, \gamma_4 \in (0,1)$ which are embedded in the *yterm* expressions determine whether the respondent assumes they must continue to pay the cost of the program when they are sick or when they are dead. We need the present discounted expected value of this stream of payments over the individual's remaining lifespan. The expected value is taken across the uncertainty about whether the individual will suffer the illness profile in question, and can be found by taking:

$$\begin{aligned} PDV\left(E\left[c_i^A *\right]\right) &= \left\{\left(1-\Pi^{AS}\right) pdvc + \Pi^{AS}\left(pdve + \gamma_3 pdvi + pdvr + \gamma_4 pdvl\right)\right\} c_i^A * \\ &= \left\{cterm_i^A\right\} c_i^A * \end{aligned} \tag{0.16}$$

The first term in the braces in the first line of equation (0.16) is the chance of staying healthy under the program, times the present discounted number of years left in the individual's nominal lifespan. The second term is the chance of suffering the illness, times c_i^{A*} weighted by the discounted number of years when neither sick nor dead and the sick-time and lost life-years weighted by the fraction of the cost that needs to be paid in those years, if any. The fact that the term in braces is the same as the *cterm* expression from equation (0.7) used in the formula for $c_i^A *$ is very convenient, since this aids us in simplifying the formula.

$$\begin{aligned} &PDV\left(E\left[\hat{c}_i^A\right]\right) \\ &= cterm_i^A\left[Y_i + \left(\frac{1}{\beta cterm_i^A}\right)\left(\begin{array}{l}\beta Y_i yterm1_i^A + \beta\left(\gamma_1 Y_i\right) yterm2_i^A + \beta\left(\gamma_2 Y_i\right) yterm3_i^A \\ +\left[\alpha_1 pdvi_i^A + \alpha_2 pdvr_i^A + \alpha_3 pdvl_i^A\right]\Delta\Pi_i^{AS} + \varepsilon_i^A\end{array}\right)\right] \\ &= Y_i\left(cterm_i^A + yterm1_i^A + \gamma_1 yterm2_i^A + \gamma_2 yterm3_i^A\right) \\ &\quad + \frac{\alpha_1}{\beta}\Delta\Pi_i^{AS} pdvi_i^A + \frac{\alpha_2}{\beta}\Delta\Pi_i^{AS} pdvr_i^A + \frac{\alpha_3}{\beta}\Delta\Pi_i^{AS} pdvl_i^A + \frac{\varepsilon_i^A}{\beta} \end{aligned} \tag{0.17}$$

In the special case where the individual assumes that they will continue to earn their regular income while sick, but not after their death, and that they will not have to pay for the program while they or sick or after their death, it will be the case that $\gamma_1 = 1$ and $\gamma_2 = \gamma_3 = \gamma_4 = 0$. In this case, the present discounted expected value simplifies to:

$$\begin{aligned} PDV\left(E\left[\hat{c}_i^A\right]\right) &= -Y_i\left[\Delta\Pi_i^{AS}\left(pdvi_i^A + pdvl_i^A\right)\right] \\ &\quad + \frac{\alpha_1}{\beta}\Delta\Pi_i^{AS} pdvi_i^A + \frac{\alpha_2}{\beta}\Delta\Pi_i^{AS} pdvr_i^A + \frac{\alpha_3}{\beta}\Delta\Pi_i^{AS} pdvl_i^A + \frac{\varepsilon_i^A}{\beta} \end{aligned} \tag{0.18}$$

In the further special case of an illness profile that involves no sick-years, just "sudden death in the current period," $pdvi_i^A + pdvl_i^A = pdvl_i^A$ and $pdvi_i^A = pdvr_i^A = 0$ so this expression will simplify to:

$$\begin{aligned} PDV\left(E\left[\hat{c}_i^A\right]\right) &= -Y_i\left[\Delta\Pi_i^{AS}\left(pdvl_i^A\right)\right] + \frac{\alpha_3}{\beta}\Delta\Pi_i^{AS}\,pdvl_i^A + \frac{\varepsilon_i^A}{\beta} \\ &= \Delta\Pi_i^{AS}\left\{\left(-Y + \frac{\alpha_3}{\beta}\right)pdvl_i^A + \frac{\varepsilon_i^A}{\beta\Delta\Pi_i^{AS}}\right\} \end{aligned} \tag{0.19}$$

If the error is assumed to be zero, so that the last term disappears, this willingness to pay measure will be proportional to the size of the risk reduction, and will not depend on the initial risk of suffering the illness profile. Furthermore, marginal and average WTP for changes in risk will be constant and equal, which is a very convenient result. Both marginal and average WTP per unit of risk reduced will be equal to the present value of annual lost income ($-Y$) and the *disutility* from the anticipation of being dead over those years, rather than alive (α_1—a negatively valued parameter), converted via the marginal utility of net income (β) to an equivalent loss in annual *value* from the anticipation of being dead in each of those years.

Unfortunately, linearity of indirect utility in net income is not supported by the data. If the indirect utility is a nonlinear function of net income, so that there is diminishing marginal utility of net income, the tidy result achieved in equation (0.19) no longer holds. In general, WTP will depend upon the initial risk level as well as the size of the risk change.

5.1.5.2 Normalization of WTP on a "statistical life"

The next step is to normalize the *WTP* amount on some arbitrary-sized risk reduction. In the literature on the "Value of a Statistical Life" (*VSL*) the convention is to normalize *WTP* on a cross-sectionally cumulative 1.00 risk change, which involves scaling up the *WTP* estimate proportionately . For the simple linear case described in the last section, rearranging equation (0.19) by dividing both sides by the size of the risk change conferred by Program A produces a convenient and simple form because in the case where utility is linear in net income, total WTP for a risk reduction is simply proportional to that risk reduction, so that average WTP per unit of risk reduction, and marginal WTP for an additional unit of risk reduction, are equal and constant:

$$\frac{PDV\left(E\left[\hat{c}_i^A\right]\right)}{\Delta\Pi_i^{AS}} = -Y_i\left(pdvl_i^A\right) + \frac{\alpha_1}{\beta}\,pdvi_i^A + \frac{\alpha_2}{\beta}\,pdvr_i^A + \frac{\alpha_3}{\beta}\,pdvl_i^A + \frac{\varepsilon_i^A}{\beta\Delta\Pi_i^{AS}} \tag{0.20}$$

This risk change, however, is a negative number. If we wish to think in terms of a positive-sized risk *reduction*, of size $\left|\Delta\Pi_i^{AS}\right|$, we could divide through, instead, by the absolute value of the risk reduction $\left|\Delta\Pi_i^{AS}\right|$, yielding the alternative formula where all terms on the right-hand side will have the opposite sign:

$$\frac{PDV\left(E\left[\hat{c}_i^A\right]\right)}{\left|\Delta\Pi_i^{AS}\right|} = Y_i\, pdvl_i^A - \frac{\alpha_1}{\beta}\, pdvi_i^A - \frac{\alpha_2}{\beta}\, pdvr_i^A - \frac{\alpha_3}{\beta}\, pdvl_i^A - \frac{\varepsilon_i^A}{\beta\left|\Delta\Pi_i^{AS}\right|} \qquad (0.21)$$

Bear in mind that the parameters α_1, α_2, and α_3 are expected to be negative, since a greater number of discounted years in each of these adverse health states is expected to decrease the individual's utility level. Thus for a linear-in-income version of the model in equation (0.13), individuals can be expected to be willing to pay a greater amount to avoid a particular illness profile, the greater their income and the greater the number of discounted years in adverse health states.

We seek only to describe the *expected* average WTP for a risk reduction for a given illness profile, by a particular type of individual (rather than to predict any individual value). Thus we set the error term ε_i to its expected value of zero and ignore the scale of its variance, given in the denominator of the last term in equation (0.21). Technically, if the transformation $f^{-1}(\cdot)$ is nonlinear, a decision to ignore the error term means that the resulting amount is the conditional median of the implied distribution of WTP, rather than its conditional mean, but we nevertheless calculate the average across all simulations of these medians.

The formula in equation (0.21) will produce something analogous to the value of a statistical life (*VSL*), which can be expected to be on the order of millions of dollars for illness profiles comparable to sudden death in the current period.

5.1.5.3 Normalization of WTP on a "microrisk reduction"

The general public (and even some uninitiated economists) experience considerable difficulty in contemplating the "value" of a "statistical life." As a consequence, in our main paper, we follow the rationale explained in Cameron (2010) and argue for normalization of WTP on a micro-risk reduction. This is achieved by dividing through not by the absolute size of the risk change, but by this risk change normalized on 0.000001.

$$WTP_{microrisk} = \frac{PDV\left(E\left[\hat{c}_i^A\right]\right)}{\left|\Delta\Pi_i^{AS}\right| / 0.000001} = \frac{PDV\left(E\left[\hat{c}_i^A\right]\right)}{\left|\Delta\Pi_i^{AS}\right|}(0.000001) \qquad (0.22)$$

5.1.5.4 The Box-Cox transformation for net income

The linear form of the function $f(\cdot)$ within the indirect utility function is rejected by our data, so we explore alternative specifications which allow for diminishing marginal utility of net income (i.e. financial risk aversion). For a Box-Cox transformation of net income, the relevant formula for the maximum annual willingness to pay is:

$$c_i^j * = Y - \left(1 + \lambda \left\{ \frac{-1}{cterm_i^j} \left[\begin{array}{l} \beta \left(\frac{Y_i^{\lambda} - 1}{\lambda} \right) yterm1_i^j + \\ \beta \left(\frac{(\gamma_1 Y_i)^{\lambda} - 1}{\lambda} \right) yterm2_i^j + \\ \beta \left(\frac{(\gamma_2 Y_i)^{\lambda} - 1}{\lambda} \right) yterm3_i^j \end{array} \right] + \alpha term_i^j \; \Delta\Pi^{jS} + \varepsilon_i^j \right\} \right)^{1/\lambda} \tag{0.23}$$

This Box-Cox transformation of net income is the function used in the main paper, for a value of the transformation parameter $\lambda = 0.45$.[32] (This value was determined via a line-search. To estimate λ simultaneously with the other indirect utility parameters would require original programming in generalized nonlinear optimization software. There is a premium on forms that lend themselves to a linear-in-parameters "index" for the estimating specification so that packaged software can be used.) By following steps analogous to those used in the linear-in-income case in the last section, it is straightforward (if a little more tedious) to arrive at formulas for the *WTP* for a microrisk reduction in this case. However, since utility is diminishing in net income, average and marginal WTP will depend upon the original risk level and the size of the risk reduction. Since WTP is no longer proportional to $\Delta\Pi_i^{AS}$, average WTP will not be constant and the original risk and the risk reduction will need to be specified in simulations of the distribution of average WTP.

5.1.5.5 The shifted logarithmic transformation for health states

In the main paper, we determined early in our analysis that the portion of equation (0.23) in square brackets that characterizes the illness profile (i.e. the discounted years in each adverse health state) is too restrictive. The data support a nonlinear specification with utility diminishing in discounted health-state years. Between a linear form and the shifted logarithmic transformation, the latter is more appropriate for these data.

In addition to switching to the logarithmic transformation, we explored a full set of second-order terms. The higher-order terms in lost life-years were robustly significant, as was an interaction term between sick-time and lost life-years. Thus we retain these terms.

Finally, since there is considerable *heterogeneity by age* in the types of illness profiles our respondents were invited to consider, it is important to control for age in these specifications.

We thus replace $\alpha term_i^j$ in equation (0.23) with:

[32] In earlier specifications we did not include scenario correction terms for cases where respondents indicate specifically that they would never benefit from the program in question, or where we there is a difference between the individual's subjective life expectancy and the life expectancies used in the choice scenarios. Without these scenario adjustment terms added, the log-likelihood-maximizing value of the Box-Cox parameter is 0.42, although the differences in the log-likelihood are small across specifications which differ only with respect to this assumed parameter value.

$$
\begin{bmatrix}
\alpha_{10}\log\left(pdvi_i^j+1\right)+\alpha_{20}\log\left(pdvr_i^j+1\right)+\alpha_{21}age_{i0}\log\left(pdvr_i^j+1\right) \\
+\alpha_{30}\log\left(pdvl_i^j+1\right)+\alpha_{31}age_{i0}\log\left(pdvl_i^j+1\right)+\alpha_{32}age_{i0}^2\log\left(pdvl_i^j+1\right) \\
+\alpha_{40}\left[\log\left(pdvl_i^j+1\right)\right]^2+\alpha_{41}age_{i0}\left[\log\left(pdvl_i^j+1\right)\right]^2+\alpha_{42}age_{i0}^2\left[\log\left(pdvl_i^j+1\right)\right]^2 \\
+\alpha_{40}\left[\log\left(pdvi_i^j+1\right)\cdot\log\left(pdvl_i^j+1\right)\right]+\alpha_{41}age_{i0}\left[\log\left(pdvi_i^j+1\right)\cdot\log\left(pdvl_i^j+1\right)\right] \\
+\alpha_{42}age_{i0}^2\left[\log\left(pdvi_i^j+1\right)\cdot\log\left(pdvl_i^j+1\right)\right]
\end{bmatrix} \quad (0.24)
$$

There is an additional interaction term involving the first sick-years term, not shown in equation (0.24), to help correct for selection bias in the parameter estimates. This selectivity-related term is described in Section 4.6. In some models, we also allow any of these α parameters, as necessary, to vary systematically with the extent to which respondent perceptions of their risks or their anticipated remaining life, for example, depart from what is asserted in their survey instrument. These terms accommodate what we call "scenario adjustment."

5.1.6 Simulated distributions for *WTP*

After estimating the parameters of equation (0.13) using maximum likelihood methods for discrete choice (discussed in the section below on Estimation), the point estimates for each of the parameters can be substituted into the formula for *WTP* for a microrisk reduction to yield a fitted *WTP* for each program offered to each individual. However, the range of programs used in our stated preference survey instrument was designed to span many of the types of health risks in the real world, the *distribution* of these illness profiles does not represent the *distribution* of health risks in the real world. These stylized and hypothetical health risks and the hypothetical programs proposed for reducing these risks are essential to the task of learning about consumer preferences, but that is the limit of our interest in them. Once we have estimates for the preference parameters, we are interested in applying them to "real" illness profiles.

The main paper outlines how the model could be employed with the range of illness profiles and the types of affected individuals in a real policy context. For the initial set of papers from this project, however, we pick just a handful of specific illustrative cases, each with a single specified illness profile that affects a single specified person. Point estimates for the *WTP* for a microrisk reduction for an illness profile could be obtained by substituting the point estimates of the indirect utility parameters into the appropriate formula for *WTP*.

However, simple use of point estimates would ignore the fact that the point estimates are not the true parameters, just *estimates* of those parameters (which are random variables). A better picture of the predictive capability of the estimated model can be obtained by simulating a distribution for the *WTP* amount, where the distribution stems from the joint density of the estimated model parameters.

We simulate *WTP* amounts by taking 1000 random draws from the asymptotic joint normal distribution of the maximum likelihood parameter estimates.[33] For each draw, we calculate the corresponding value of the *WTP* for a microrisk reduction. Across all 1000 draws, we build a "sampling distribution" for the *WTP* amount that reflects the variability in the estimated indirect utility parameters. In the tables in our papers, unless indicated otherwise, we

[33] The "drawnorm" utility in Stata is very useful for this type of exercise.

report the empirical mean, and the empirical 5th and 95th percentiles of this distribution of 1000 values. This information provides a sense of the central tendency and the dispersion of the quantity of interest: *WTP* for a risk reduction.

As always, when simulating such distributions, it is readily apparent that there is a tradeoff between bias and efficiency. While it may be tempting to include less statistically significant explanatory variables in the choice model, the presence of insignificant coefficients can inflate the variance-covariance matrix for the parameters and this can result in very wide 90 percent intervals for the simulated distribution of *WTP* for a microrisk reduction. We typically worry that omission of insignificant variables may incur some degree of bias in the estimates of the remaining indirect utility parameters. However, given the extent of the randomization of the illness profiles and program costs in this study, there is little concern about omitted variables bias. Parsimonious specifications are likely to be appropriate.

5.2 Estimation

Now we provide some background concerning the estimators used to produce the vector of maximum likelihood parameter point estimates and the parameter asymptotic variance-covariance matrix reported in the paper. This discussion assumes that the reader is familiar with conventional textbook treatments of models for unordered multiple discrete choice, for example, as covered in section 23.11 of Greene (2008).

The choice sets faced by each respondent on each choice occasion in our study consist of three alternatives: Program A, Program B, or Neither Program. The "dependent variable" in this context is actually a set of three indicator variables, switched "on" if the corresponding alternative is chosen, and "off" if it is not. The explanatory variables all differ across alternatives: net income will depend upon which alternative is chosen because each program has a different cost. The chance of suffering each featured illness profile, interacted with the nature of that illness profile, will also differ across all three alternatives. Thus each explanatory variable differs both across individuals and across alternatives within each choice set the individual faces, so the "conditional" logit model is appropriate, as in section 23.11.2, p. 846-847 of Greene (2008). Our constructed explanatory variables, used in our estimating specifications, are examples of Greene's x_{ij} variables.

5.2.1 Panel data: Fixed Effects?

The first notable thing about the structure of our data on respondents' three-way multiple discrete choices is that these are effectively "panel" data. Each respondent, typically, provides us with five different choices. With panel data, there is always a question whether a set of slope coefficients, estimated using simply the pooled data without recognition of its panel nature, might be affected by heterogeneity bias. (Heterogeneity bias is a form of omitted variables bias, where the explicit explanatory variables are correlated with unobserved forms of heterogeneity across individuals, so that the estimated coefficients are biased). Fortunately, the randomized design of all of our choice sets, conditional only on the age and gender of the respondent and the plausibility of some types of outcomes, means that the x_{ij} variables in our models are unlikely to be correlated with any omitted variables, especially since we control for the respondent current age in our models.

Nevertheless, the fact that we have repeated choices for each person in our sample immediately led us (and almost every reviewer of our work) to a concern that appropriate panel-oriented econometric methods should be used with these data. The parameters of our model are thus estimated using the fixed effects conditional logit choice model as implemented in the Stata 10 econometric software package. The model is described in considerable detail in the Stata 10 Reference Manual under the heading "clogit – Conditional (fixed effects) logistic regression" (p. 285-287).

5.2.2 Panel data: Fixed or random parameters?

A further possibility is that our choice models should be estimated using random-parameters logit models. These model permits each utility parameter to be individual-specific and the same across all five choices made by the any single individual. However, it assumes that these individual-specific parameters are a random draw from a joint distribution of utility parameters in the population. The goal is to estimate both a central tendency and a dispersion for each utility parameter, to allow explicitly for unobserved forms of heterogeneity in preferences.

In the next section, we describe first the biostatistical version of the fixed-effects logit model, as summarized in the documentation for Stata's algorithms. We then provide an alternative description of these models, from an econometric perspective, as explained in Greene (2008). In Section 4, we describe the results of using random parameters (mixed) logit models with our data, along with our rationale for preferring to estimate systematically varying parameters, rather than randomly varying parameters.

5.3 Fixed effects versus no fixed effects

Breslow and Day (1980), pages 247-279, Collett (2003), pages 251-267, and Hosmer and Lemesow (2000), pages 223-259, provide the biostatistics version of "conditional logistic regression." Hamerle and Ronning (1995) also describe the fixed-effects logit, but Chamberlain (1980) is the standard econometrics reference for this model. We provide both the biostatistical and the econometric perspectives on this model in the two sections to follow:

5.3.1 Biostatistical Perspective

We use the pre-programmed algorithms in the Stata software package to estimate our fixed effects logit models. Stata's description of the estimator is couched in terms of the biostatisticial approach to these models. For those who are most familiar with that approach, we adapt the description in the Stata manual, tailoring it to the application of the model in this paper, and using conformable notation, let $i = 1,...,n$ denote respondents and let $k = 1,...,5$ denote the five choice scenarios presented to each respondent. We will start with an exposition which assumes just the choice between "Program A" and "No program" (N). Let y_{ik} be the dependent variable taking on values 1 if the program is chosen and 0 if no program is chosen. Let $y_i = (y_{i1},...,y_{i5})$ be the outcomes for the i^{th} respondent. Let x_{ik} be a row vector of covariates (i.e. the explanatory variables listed as regressors for our choice models). Let

$$h_{1i} = \sum_{k=1}^{5} y_{ik} \tag{0.25}$$

be the observed number of ones for the dependent variable for the i^{th} respondent. In the biostatistical version of the model, practitioners would say that there are h_{1i} "cases" matched to $h_{2i} = 5 - h_{1i}$ "controls" for the i^{th} respondent.

In the analysis, we consider the probability of a possible value of y_i, the vector of outcomes, conditional on $\sum_{k=1}^{5} y_{ik} = h_{1i}$ (Hamerle and Ronning, 1995, equation 8.33; Hosmer and Lemeshow, 2000, equation 7.4),

$$\Pr\left(y_i \mid \sum_{k=1}^{5} y_{ik} = h_{1i}\right) = \frac{\exp\left(\sum_{k=1}^{5} y_{ik} x_{ik}\beta\right)}{\sum_{d_i \in S_i} \exp\left(\sum_{k=1}^{5} d_{ik} x_{ik}\beta\right)} \tag{0.26}$$

where d_{ik} is equal to 0 or 1 with $\sum_{k=1}^{5} d_{ik} = h_{1i}$ and S_i is the set of all possible combinations of h_{1i} ones and h_{2i} zeros. There are $\begin{pmatrix} 5 \\ h_{1i} \end{pmatrix}$ such combinations, but we fortunately do not need to count all these combinations to compute the denominator in equation (0.26), since it can be computed recursively. Denote the needed denominator as:

$$g_i\left(5, h_{1i}\right) = \sum_{d_i \in S_i} \exp\left(\sum_{k=1}^{5} d_{ik} x_{ik}\beta\right) \tag{0.27}$$

Consider, computationally, how g_i changes as we go from a total of 1 choice set per person to 2 choice sets, and so on. Doing this, we derive the recursive formula:

$$g_i\left(5, h\right) = g_i\left(4, h\right) + g_i\left(4, h-1\right)\exp\left(x_{i5}\beta\right) \tag{0.28}$$

where we define $g_i\left(5, h\right) = 0$ if $5 < h$ and $g_i\left(5, 0\right) = 1$.

The conditional log-likelihood for this problem is:

$$\ln L = \sum_{i=1}^{n} \left\{ \sum_{k=1}^{5} y_{ik} x_{ik}\beta - \log g_i\left(5, h_{1i}\right) \right\} \tag{0.29}$$

where the derivatives of the conditional log-likelihood can also be computed recursively by taking derivatives of the recursive formula for g_i.

The documentation for Stata 10 indicates that computation time is roughly proportional to $p^2 \sum_{i=1}^{n} 5 \min\left(h_{1i}, h_{2i}\right)$, where p is the number of independent variables in the model. If $\min\left(h_{1i}, h_{2i}\right)$ is small, computation time is not an issue.

5.3.2 Econometric Perspective

Based on Greene (2008), Ch. 23.5.2, and the references cited therein, we can adapt the discussion of the choice probabilities employed in Chamberlain's conditional likelihood function to a simple case which conveys the intuition of the fixed effects logit approach. Suppose the

systematic portion of the indirect utility differences associated with each individual include a component that is constant for any one individual but differs across individuals: $\alpha_i + x_{ik}\beta$. In the context of the models explored in our age-differentiated translog-type specifications, the x_{ik} vector consists of the thirteen basic explanatory variables constructed from our raw data, where the model involves five choice scenarios per person, each concerning three alternatives. To keep the algebra simple, consider the binary choice case, rather than the three-way choice case, with the recognition that it can be generalized to the three-alternative case considered in the body of our paper.

The unconditional likelihood function, when there are K choices involving just a pair of alternatives for each individual, will take the following form, where the regressors, x_{ik}, are implicitly the *differences* between the attributes of the two alternatives between the "1" and the "0" outcome (often the status quo outcome for which attribute levels are normalized to zero):

$$L = \prod_{i=1}^{n}\prod_{k=1}^{K}\left(\frac{\exp\left(\alpha_i + x_{ik}\beta\right)}{1+\exp\left(\alpha_i + x_{ik}\beta\right)}\right)^{y_{it}}\left(\frac{1}{1+\exp\left(\alpha_i + x_{ik}\beta\right)}\right)^{1-y_{it}} \tag{0.30}$$

For the five different three-way choices made by respondents in our study, the corresponding unconditional likelihood function would involve three distinct indicators, y_{Aik}, y_{Bik}, and y_{Nik} that take the value 1 if the corresponding alternative among A, B, and N is chosen, and the value 0 otherwise. The regressors are the attributes associated with each alternative, normalized on their levels for the "Neither Program" alternative to permit estimation of a unique parameter vector. This would be an ordinary pooled-data conditional logit model except for the individual-specific constant which shifts the systematic utility associated with every non-numeraire alternative.

$$L = \prod_{i=1}^{n}\prod_{k=1}^{5}\left[\begin{array}{l}\left(\dfrac{\exp\left(\alpha_i + \left(x_{Aik} - x_{Nik}\right)\beta\right)}{\exp\left(\alpha_i + \left(x_{Aik} - x_{Nik}\right)\beta\right)+\exp\left(\alpha_i + \left(x_{Bik} - x_{Nik}\right)\beta\right)+1}\right)^{y_{Aik}} \\ \left(\dfrac{\exp\left(\alpha_i + x_{iA}\beta\right)}{\exp\left(\alpha_i + \left(x_{Aik} - x_{Nik}\right)\beta\right)+\exp\left(\alpha_i + \left(x_{Bik} - x_{Nik}\right)\beta\right)+1}\right)^{y_{Bik}} \\ \left(\dfrac{1}{\exp\left(\alpha_i + \left(x_{Aik} - x_{Nik}\right)\beta\right)+\exp\left(\alpha_i + \left(x_{Bik} - x_{Nik}\right)\beta\right)+1}\right)^{y_{Nik}}\end{array}\right] \tag{0.31}$$

Given the fundamental nonlinearity of the model, we cannot just use differences from within-group means (as we might do in a least-squares context) to sweep out the "intercept" values in the logit "index", $\alpha_i + \left(x_{jik} - x_{Nik}\right)\beta$. Instead we use a clever insight from Chamberlain. His approach relies on the sequences of choices observed in the set of choices for each person. Suppose there are just two choices for each person, as in equation (0.30) (multiple-alternatives and several choice occasions just mean messier algebra). Then the person could choose (1,1), (1,0), (0,1) or (0,0). Chamberlain conditioned the probability of a particular pattern of choices on the outcome that the sum of the indicators took on each particular value.

The conditional likelihood is given by:

$$L^c = \prod_{i=1}^{n} \Pr\left(Y_{i1} = y_{i1}, Y_{i2} = y_{i2} \left| \sum_{k=1}^{K} y_{ik} \right. \right) \tag{0.32}$$

For example, the probability that the pair of choices will be (0,1) when the sum of the indicators is one is given by:

$$\Pr(0,1 \mid sum = 1) = \frac{\Pr(0,1)}{\Pr(0,1) + \Pr(1,0)} \tag{0.33}$$

The probability can be built from the binary probit probabilities in each of the two choice occasions. To simplify the notation, assume this is a case where the levels of the attributes have been normed on the status quo alternative, so that $\alpha_i + \left(x_{1ik} - x_{0ik} \right) \beta$ can be written simply as $\alpha_i + x_{ik}\beta$ with the regressors understood to be the difference in attribute levels between the "1" and the "0" alternatives. The error terms are still assumed to be uncorrelated, so the key insight is that each of the joint probabilities for the pairs of outcomes in the numerator and denominator of equation (0.33) can be written as the product of the probabilities of each outcome by itself:

$$\Pr(0,1 \mid sum = 1) = \frac{\left[\frac{1}{1+\exp\left(\alpha_i + x_{i1}\beta\right)} \cdot \frac{\exp\left(\alpha_i + x_{i2}\beta\right)}{1+\exp\left(\alpha_i + x_{i2}\beta\right)} \right]}{\left[\frac{1}{1+\exp\left(\alpha_i + x_{i1}\beta\right)} \cdot \frac{\exp\left(\alpha_i + x_{i2}\beta\right)}{1+\exp\left(\alpha_i + x_{i2}\beta\right)} \right] + \left[\frac{\exp\left(\alpha_i + x_{i1}\beta\right)}{1+\exp\left(\alpha_i + x_{i1}\beta\right)} \cdot \frac{1}{1+\exp\left(\alpha_i + x_{i2}\beta\right)} \right]} \tag{0.34}$$

We can now see why this conditional likelihood is attractive…the denominator terms in the expressions above and below the line will cancel, leaving just:

$$\begin{aligned} \Pr(0,1 \mid sum = 1) &= \frac{\exp\left(\alpha_i + x_{i2}\beta\right)}{\exp\left(\alpha_i + x_{i2}\beta\right) + \exp\left(\alpha_i + x_{i1}\beta\right)} \\ &= \frac{\exp\left(\alpha_i\right)\exp\left(x_{i2}\beta\right)}{\exp\left(\alpha_i\right)\left[\exp\left(x_{i2}\beta\right) + \exp\left(x_{i1}\beta\right)\right]} \\ &= \frac{\exp\left(x_{i2}\beta\right)}{\exp\left(x_{i1}\beta\right) + \exp\left(x_{i2}\beta\right)} \end{aligned} \tag{0.35}$$

The other way to get a sum of 1 will have a probability that is just the complement of this probability, with $\exp\left(x_{i1}\beta\right)$ in the numerator instead.

In the case of two binary choices, there are just three possible sums: one way to get a sum of 2; two ways to get a sum of 1, and one way to get a sum of 0. Notice that somebody who chooses "all zeros" or "all ones" will yield a sum of zero or a sum of *K* (the number of choices, here just two). Since there is only one way to do each of these things, people who always choose the same alternative will have a conditional probability of one, and the log of one is zero, so they

add nothing to the log of the conditional likelihood. Their choices will not contribute to the estimation of the slope parameters in the vector β. Only the cases with sums between 2 and *K-1* are helpful.

When the objective function is constructed from these types of *conditional* probabilities, we allow for individual-specific "lumps" of utility in the amount of α_i for each respondent, although we forgo the ability actually to estimates these parameters (as we do in a fixed effects model in a least-squares context when the slope coefficients are estimated using the method of deviations from within-group means). However, the slope coefficients, β (here interpreted as marginal utility parameters associated with each attribute) are estimated assuming the existence of heterogeneity in the α parameters.

5.3.3 Hausman test for fixed effects

To test a fixed effects logit against an ordinary logit, we normally use a Hausman-type test concerning what happens to the vector of slope coefficients across the two specifications. If preferences are homogeneous (i.e. if there is no need for the fixed effects model) both the ML and the CML are consistent, but the Chamberlain estimator is inefficient (because it will not really use the information from people who chose the same alternative on all of their choice occasions, and it does not take advantage of the constraint that $\alpha_i = \alpha$). The Hausman test is:

$$\left(\hat{\beta}_{CML} - \hat{\beta}_{ML}\right)'\left(Var\left[CML\right] - Var\left[ML\right]\right)^{-1}\left(\hat{\beta}_{CML} - \hat{\beta}_{ML}\right) \sim \chi^2(k) \qquad (0.36)$$

where k is the number of slope parameters (e.g. marginal utilities in an additive RUM model). A large value of this test statistic says that moving to a fixed effects model has made a big enough difference in the slopes for us to believe that the homogeneous model was too restrictive.

Fixed effects logit models can be invoked in Stata by using the command:

clogit best x1 x2 …, group(personid);

where the data have been entered with one row for each alternative, three rows for each choice set, and k choice sets (typically five) per person. The variable "best" is a binary indicator for the chosen alternative in each choice set, and the majority of respondents in the sample will each account for fifteen rows (5 x 3) in the data. The x variables are the explanatory variables, both individual- and alternative-specific, which we use to account for respondents' choices.

For our preferred specification, the results of the Hausman test are shown in Figure 5-2. Notice that the differences in the estimated parameters are relatively minor and that the calculated χ^2 test value rejects the null hypothesis only at the 13% level, although the algorithm is hampered by the fact that the difference between the parameter variance-covariance matrices for the two models is not positive definite (where the difficulty concerns the term in discounted recovered/remission years, a variable which is individually statistically significantly different from zero only at the 10% level in both the fixed effects model and the non-fixed effects model).

There is little a priori reason to anticipate that a fixed effects specification will be necessary because the levels of all of the main regressors have been assigned randomly across choice sets and across individuals. The only source of concern will stem from the appearance of the interaction terms in age and age-squared which shift three of the basic coefficients, and the

selection correction interaction term involving each individual's fitted probability of participating in the estimating sample (relative to the original 525,000 recruiting contacts for Knowledge Networks). While the results of the Hausman test reproduced in Figure 5-2 suggest that there is no strong evidence of the need for a fixed effects conditional logit model, we employ it for its greater generality. Failure to exploit the panel dimension of our data would invite the criticism that we have somehow obscured relevant unobserved heterogeneity in consumer preferences.

5.4 The scale factor (heteroscedasticity in the errors?)

In our basic models, we assume that the error term in our model is homoscedastic. Of course, much interest in recent years has been focused on the possibility that the "scale factor" differs according to the characteristics of the individual or the context of the choice. It is of course entirely possible to define variance components, unique to each choice set (or each risk-reduction program), according to which diseases are included among the two which are mentioned in each set.

We have explored models with the error variance normalized to unity for heart disease, and separate multiplicative terms for each of the eleven other illness labels, switched on or off according to whether an illness profile bearing that label is involved in each choice. Estimation of this model, of course, means shifting to general nonlinear function-optimizing software (we use Matlab). To keep the parameter space manageable as we conducted preliminary explorations of the need for heteroscedastic errors, we used a simple five-parameter specification that employs a quadratic form in net income and the three health-state duration variables entered in linear form, rather than logarithmic form. If there is mischief in the error term, it is often most pronounced when the systematic portion of the model is underspecified in some way. These models converged readily.

Expanding the model to include eleven extra disease-specific error-term dispersion parameters (relative to that for the numeraire illness), the maximized value of the log likelihood improves by less than eight points, which suggests that the heteroscedasticity parameters are not jointly significant. Only one individual parameter comes remotely close to statistical significance (i.e. the coefficient for breast cancer has an asymptotic t-test statistic of -1.59). The coefficients of the logit index for the heteroscedastic model average about 1.27 times the magnitude of the coefficients from the homoscedastic model, as shown in Table 5-3. Although not shown in this table, the two sets of parameter estimates have overlapping confidence intervals.

Of course, logit coefficients are known only up to a scale factor. The coefficients for the homoscedastic model are normalized upon the assumed common error dispersion shared by all types of illnesses. For the heteroskedastic model, the logit coefficients are normalized on the error dispersion for the omitted category of illness, heart disease, which will likely be different from the "average" dispersion across all illness types. It is not surprising that the coefficients of the logit index for the heteroskedastic model are about 1.27 as large as those in the homoscedastic model. This would be consistent with the error dispersion related to heart disease being about 0.79 times the average dispersion in the homoscedastic model, which could easily be the case.

We also attempted a heteroskedastic model that included additional distinct shifters on the error dispersion terms related to the individual's subjective risk of each type of disease (rated on a -2 to +2 scale). Unfortunately, this model with 11+12=23 dispersion shifters could not be

coaxed to convergence. Explicit accommodation of a lot of systematic heterogeneity can often mitigate heteroscedasticity that shows up in simpler models. We have used these same data to estimate models that allow for heterogeneity not only by age, but by illness labels, subjective illness risks, comorbidity, household structure, and other factors, and the main results are highly consistent. If no significant heteroscedasticity by disease type shows up in this relatively simple model, we believe that the odds are probably smaller that it might cause significant distortions of the primary inferences in the much more general models that we explore in our various papers.

5.5 Random-parameters logit models

Over the last decade, it has become increasingly easy to consider random-parameters variants of multiple discrete choice models. The familiar "mixed logit" model (e.g. Greene (2008), p. 851-859) is an important alternative specification to consider in this application. Using Kenneth Train's mxlmsl.m Matlab algorithm (mixed logit by maximum simulated likelihood), we have estimated mixed logit models assuming normal distributions for all thirteen basic marginal-utility-related coefficients featured in our age-differentiated Translog-type specifications.

A key insight from this exercise concerns the distribution of ages in the estimating sample. This distribution is depicted in Figure 5.3. If we allow for normally distributed coefficients in a random parameters model, all of the previously identified systematic variation due to age heterogeneity in our basic model (the one that allows linear or quadratic shifters in age on several of the marginal utility parameters) is absorbed instead by the random parameters. Furthermore, only the four basic parameters, β_0 (on the net income term), and α_{10}, α_{20}, and α_{30} (on each of the discounted future health state terms) display statistically significant heterogeneity (in terms of the estimated dispersion in the random parameter).

5.5.1 Results: Random parameters specifications

If we generalize our basic specification to allow for interaction terms in age and age-squared also to shift the estimated β_0 parameter, neither of these interaction terms bears a statistically significant coefficient, so besides age there is apparently some other source of unobserved heterogeneity in this marginal-utility-of-income parameter. (Other candidate sources of systematic variation in the marginal utility of income are discussed in Section 5.6.11). However, it seems clear that the age variable is a prominent contributor to heterogeneity in the slope coefficients which capture the marginal (dis)utility of future adverse health states. This heterogeneity with respect to age is a key consideration in any model of health risk reduction preferences, so we expressly *do not* wish to subsume it with all other unspecified sources of heterogeneity in a mixed logit model. Thus we opt for a conventional non-random parameters specification in this application. In other research using these data, we explore for other possible dimensions of heterogeneity in the marginal utility of income parameter, notably in our "comorbidity" paper, but those analyses are beyond the scope of the main paper.

5.6 Alternate Specifications

5.6.1 Preliminary models

In Table 5-4, we consider first the implications of our data in the context of the simplest *ad hoc* specification. Model 1 reveals that the two main features of each program we describe in our choice scenarios—namely, its cost and the size of the risk reduction it would achieve—are both strongly statistically significant determinants of people's choices. We then show, in Model 2, that the two most important features of each *illness profile*—namely, the prospective sick-years and lost life-years—are also strongly statistically significant in explaining choices. In the final specification in this table, Model 3, we implement the four-parameter structural model outline in Section 5.1.5.1, imposing a Box-Cox transformation with $\lambda = 0.42$ as the function $f(\cdot)$.[34]

Model 3 is a homogenous-preferences specification, estimated without sign restrictions, and shows robust significance and the expected signs on all four primary parameters. The estimated marginal utility of income is positive and declines with the level of income. The marginal utilities of discounted prospective sick-years, post-illness recovered/remission-years, and lost life-years are all negative and very strongly significantly different from zero. Simple intuition might suggest that death should be perceived as "worse" than illness and recovery/remission. However, it is important to keep in mind that the units involved are discounted single years in each health state. In many illness profiles, there are more life-years lost than there are sick-years, but the lost life-years are always further into the future, so they are discounted more heavily. Thus the marginal utility per discounted health-state year does not convey the overall disutility of *total* future time in that state. Also, the relatively large (dis)utility associated with recovered/remission state reflects the seriousness of the major illnesses our survey describes. Rightfully, respondents do not interpret being recovered or in remission from any of this list of major illnesses as being equivalent to the pre-illness "healthy" state, which would produce a zero coefficient. For example, there may be considerable anticipated disutility from the prospect of living as a cancer or heart-attack survivor, relative to the respondent's current health.[35]

In the main paper, we quickly relax the assumption that the marginal utilities from each prospective future health state are independent of the duration of that state and the durations of other health states that characterize the illness profile in question.

5.6.2 Appropriate transformation for health state durations

In the main paper, we first consider a model that is linear in the discounted prospective sick-years, recovered/remission years, and lost life-years. The parameters of this model are identical to the underlying parameters in the future-period indirect utility function. However, we find that a model which takes the present discounted time in each future health state as the relevant

[34] The curvature in the net income term allows for risk aversion with respect to financial risk. Eeckhoudt and Hammitt (2004) find that this type of risk aversion increases *WTP* for risk reductions in definable cases, but that in general, the relationship is theoretically ambiguous. We note that the structural form in Model 3 yields a somewhat poor overall fit than that attained with *ad hoc* Model 2. However, this structural form is the feature that permits us to calculate rigorously the corresponding option price that is our *WTP* measure.

[35] The evidence about the marginal (dis)utility of a discounted recovered/remission-year also does not involve diabetes or Alzheimer's disease, since it was not possible to describe credible scenarios with recovery from these diseases.

characteristics of the entire illness profile may be superior, and that a shifted logarithmic transformation seems to dominate an ordinary linear function of the present discounted durations.

It is impractical to do a four-way grid search to establish different transformations for each of the four main variables in the estimating specification. However we have constrained the Box-Cox parameter for the net income term, $f(Y)$ to be $\lambda = 0.45$ in our more-general specifications. We have also conducted a single line-search across values of and additional parameter θ, a Box-Cox transformation parameter shared by all of the discounted durations in each health state (shifted by one, to ensure that a zero duration corresponds to a zero value of the transformed variable as well). In Figure 5-4, we show that the maximized log-likelihood for the four-parameter conditional-logit model is relatively insensitive to the choice of θ for values less than zero (which would correspond to the logarithmic transformation adopted in the paper). However, the maximized log-likelihood begins to drop off at a distinctly faster rate for parameter values greater than zero.

Based on Figure 5-4, we elect to use the simple shifted logarithmic transformation for each of the discounted health-state durations in our model.

5.6.3 Correcting for scenario adjustments

In Cameron et al. (2010a) we explored some early specifications using these data and determined that it was not appropriate to ignore the information in some of the debriefing questions that were posed to respondents. We advocate the use of information about the extent to which respondent's subjective perceptions of their likely illness profiles deviate from those described in the choice tasks on the survey instrument. To demonstrate this method, we normalize on "acceptance" of the description in each choice scenario and allow the parameters of the utility function to differ systematically with the extent to which the respondent's subjective perceptions depart from the assumptions in each choice scenario.

5.6.3.1 Adjustment for "would never benefit from program"

As one type of correction, we used a dummy variable for whether the respondent stated explicitly that they would *never* benefit from the risk reduction program in question. This is the last option, for each illness, in the survey form depicted in Figure 5-5. (Thus a zero value for this variable corresponds to "acceptance" of the program's stated benefits.)

5.6.3.2 Adjustment for difference between subjective and stated life expectancy

Instead of using the overestimates of the latency as scenario adjustment variables, we resort to another correction that seems to be relatively unambiguous and should probably be made on a priori grounds. Before we introduced the choice tasks in our survey, each respondent was told his or her "nominal life expectancy" because it was necessary to frame the illness profiles in terms of how much life they might be able to live if they did not suffer each of these illnesses. As already noted earlier in this Handbook, many respondents in our pre-tests rejected actuarial life expectancies, so we made a strategic decision to overstate life expectancies by

enough years (eight) to minimize the chance of this happening. At the end of our survey, however, respondents we questioned explicitly about their life expectancies, as in Figure 5-6. About 20% of our sample still thought they would live longer than the life expectancy we had used in their choice scenarios. As Figure 5-7 reveals, however, to preclude as fully as possible scenario rejection based on the stated life expectancy being too short we need to add more than the *average* number of extra years that would have been necessary. (About two more years, rather than eight, would have matched the *average* of the subjective life expectancies.)

If the individual finds this stated life expectancy to be subjectively implausible, this mismatch may have a systematic effect on the estimated utility parameters. Thus we entertain the difference between the respondent's own subjective life expectancy and the nominal life expectancy used in our survey as another scenario adjustment variable. Since we created this disparity through the design of our choice scenarios relative to the individual's current age, it seems incumbent upon us to explore corrections for any mischief this necessary design decision may have created.

In Cameron et al. (2010b), we used a simple linear term in "respondent's subjective overestimate of life expectancy" as a scenario adjustment shifter on the indirect utility parameters. Since that paper, however, we have explored some further generalizations. It seems less restrictive to permit the effects of an *over*estimate of life expectancy to be different from the effects of an *under*estimate of life expectancy, so we use separate variables for these two effects. Furthermore, the lower tail of the empirical distribution (with a few life expectancies as much as 30 to 50 years less than used in the choice scenarios) seems questionable. Hence we temper the influence of the extreme departures by using the logarithm of the absolute value of the discrepancy as the scenario adjustment variable, entering this variable separately for positive and negative departures.

5.6.3.3 Adjustment for "did not specifically consider affordability"

Form 47 of our survey asked the respondent "Did you consider whether you could actually afford to pay for these programs over your lifetime?" Under ideal choice conditions, the respondent would have answered "yes." However, across the estimating sample, only 46.4% of respondents answered "yes." About 39.7% said only "somewhat," and 13.8% said "no." This self-reported behavior does not comply with how respondents were asked, beforehand, to consider the choices. Specifically we reminded them: "In surveys like this one, people sometimes do not fully consider their future expenses. Please think about what you would have to give up to purchase one of these programs. If you choose a program with too high a price, you may not be able to afford the program when it is offered."

Since the answer to this debriefing question reveals an "unauthorized" adjustment by respondents, it is appropriate to correct for this adjustment during estimation. We generate an indicator variable for this failure to comply, called *affordmiss*, which equals one if the answer on Form 47 is anything other than "yes." We would expect people who did not fully consider whether they could afford to pay would be less sensitive than they ought to be to the costs of the program. Their estimated coefficient on the net income term should be smaller.

It is possible to break out the *affordmiss* variable into two components, one for people who answered "some," and one for people who answered "none" (where we include refusals to respond in this latter category). However, while the point estimates on both coefficients are negative and significant, the ranking of the individual sizes of the coefficients is counterintuitive.

Since the confidence intervals overlap substantially, we opt to combine all of the non-"yes" responses into one category.

The only coefficient that is shifted to a statistically significant extent by the *affordmiss* variable, logically, is the coefficient on the *bXterm* net income variable, which is logical. In some specifications, the t-tet statistic for the coefficient on the relevant interaction term exceed 5 in absolute value. Thus it will be important to net out the effect of people who self-report paying incomplete attention to their ability to pay for the health risk reduction programs. We are interested in isolating the coefficient on the net income variable that applies for people who do pay attention to their ability to afford the goods in question. The correction makes the baseline marginal utility larger, which will lower predicted WTP.

5.6.3.4 Discontinued: adjustment for overestimate of the latency

As a second shifter, we built a variable that measured the minimum overestimate of the latency of the disease. This departure is zero if the stated latency of the illness in question falls within the interval checked by the individual in the relevant question in Figure 5-5. If they believed the benefits would start later than this, their minimum overestimate of the latency is positive. If they believed the benefits would start earlier than was stated, their minimum overestimate of the latency is negative. Employing this differential as a shifter on the utility parameters in the model was a bit more of a stretch because we required, in effect, that the respondent perceived no benefit from the program until the disease would otherwise produce at least "moderate pain and disability."

Cameron et al. (2010a) explains how this variable was constructed, and we have explored such a correction for the models in our main paper as well. However, we now believe that this debriefing question may not have been sufficiently explicit. We concede that individuals could reasonably have expected that they could benefit from the risk reduction program before "moderate pain and disability" would develop. Their quality of life could indeed be better, even prior to significant symptoms, if they participated in the program. Furthermore, if we employ these corrections, we notice as a practical matter that the models which retain the full suite of potential scenario adjustment variables produce changes to the "zero-departure" indirect utility parameters that lead to negative fitted WTP values for our "end-of-life" illness profiles. We take this as further evidence that these corrections may be inappropriate.

5.6.3.5 Not implemented : adjustment for "which shortens life most" response

Form 16 of our survey was intended to ascertain whether respondents understood the illness profiles. Despite the apparent acceptance of this form during the cognitive interviews with our pre-test subjects, we now believe that the form was poorly designed because the actual question invited confusion. We presented respondents with the two diseases in their tutorial section and the portion of the choice table that describes age at recovery (if any) as well as the effects of each disease on their life expectancy. But then we asked them "Which one shortens your life the most?" We should have asked "Which one leads to an earlier death?" It was a poor choice to use the words "shortens" and "most" in the same sentence.

We believe it was our poorly designed question that caused more than half of our respondents (51.05%) to answer this question incorrectly. Had we phrased the question better,

we probably would have expected more people to get this comprehension test right than the question about the risks.

5.6.4 Final baseline specification, other than incidental variables

The final set of twelve key illness-profile terms in our "basic" specification has been arrived at through extensive exploration of our data. As always, intuition and the underlying economic theory dictate which variables one should *expect* to have a role in explaining consumer choices. These factors dictate that there should be terms in the present discounted duration in each of the three adverse health states, as we assumed in the discussion of the linear model above. Relevant dimensions of non-linearity must be determined by an appeal to the data. In general, we choose to retain lower-order terms with marginally insignificant coefficients when the estimated coefficient on a corresponding higher-order term or interactions proves to be statistically significant. With that constraint, we have explored a number of generalizations and then backed off when they appear not to be necessary. To arrive at the specification in the expression in (0.24), we generalized all three shifted logarithms of the discounted health states by permitting their coefficients to vary systematically with the respondent's age and with age-squared. We also explored all three pairwise interactions, with their coefficients also allowed to differ with age and age-squared. However, the terms for the pairwise interaction between sick-years and recovered/remission years were persistently unhelpful in explaining respondents' choices, so these terms were dropped.

We then introduced the various scenario adjustment variables (discussed elsewhere in this document). Each of these variables was interacted with all of the basic variables, their age-related shifters, and their relevant higher-order and interaction terms. The non-scenario-adjustment variables were forced into a model that otherwise allowed stepwise elimination of irrelevant scenario adjustment variables. The baseline non-scenario-adjustment coefficients were inspected for their remaining contributions. Variables whose coefficients had t-test statistics less than one in absolute value were considered candidates for exclusion. This led us to drop both the linear and quadratic age shifters on the sick-years term, and the quadratic age shifter on the recovered/remission-years term, as well as all three basic terms in the interaction between recovered-remission years and lost life-years (the interaction and the age and age-squared shifters on its coefficient).

With these base variables pruned from the specification, stepwise methods were again used to determine which of the original universe of scenario adjustment variables remained relevant, conditional on the list of basic health state regressors in (0.24). Our final preferred specification involving this list of basic regressors involves controls for the influence of scenario adjustments on the coefficients for six of these regressors. Note that the age interactions for the sick-years term are statistically no different from zero for respondents who accept the main features of the choice scenarios, so we drop these baseline age interactions. However, these age interactions have coefficients which are statistically significant for respondents who claim that that the program in question will never benefit them, and the coefficient on the quadratic term in age is significant for respondents whose life expectancy is less than that stated in the choice scenarios. We honor this estimated heterogeneity in preferences by retaining scenario adjustment shifters on three terms which themselves do not survive in the base model shown in (0.24). In our WTP calculations, of course, we simulate an absence of any scenario adjustment, so these terms are set to zero. However, their influence on the estimated coefficients on the baseline variables remains relevant.

Table 5-5 shows the differences in indirect utility parameters as we generalize our specification from (1) a simple four-parameter model, to (2) a model with significant nonlinear terms in discounted health-state years, to (3) a model with a correction for sample selection, to (4) a model with our final working scenario adjustment corrections (for programs that will never benefit the individual and subjective over- and under-estimates of life expectancy), and (5) finally to a model where we permitted over-and under-estimates of disease latency (which we argue now is probably an inappropriate correction). Thus the parameter estimates in column 4 of this table are the results features in our main paper.

Table 5-6 shows the corresponding implications of these five different specifications for all of our simulated WTP amounts. Again, the results for column 4 of this table correspond to the estimates provided in our main paper.

5.6.5 Different assumptions about the fixed discount rate

In addition to the specifications reported in the main paper, we have explored a variety of other possible specifications. One key assumption in the estimation concerns the common discount rate attributed to all respondents.

The basic results in the main paper reflect the assumption of a 5% discount rate. However, we also report the distributions of simulated *WTP* estimates if alternative assumptions are made about this discount rate. These alternative *WTP* results are derived using the parameter estimates reported in column 4 of Table 5-5. Recall that the discount rate assumption is invoked during the construction of variables used in our estimation. When a different discount rate is assumed, different "present discounted" variables must be reconstructed based on that different assumed discount rate. Since these calculated variables will be somewhat different, so will be all of the parameter estimates produced by the model.

Table 5-9 compares the parameter estimates produced for different alternative assumptions about the common discount rate used by respondents. We consider common fixed discount rates of 3%, 5% and 7%. Table 5-10 mirrors Table 5-6 by disclosing the effects of different discount rate assumptions on the resulting estimates of each of the WTP measures considered in a basic model. In some cases, the discounting assumption makes a considerable difference to WTP estimates, but not in other cases. As is to be expected, discount rates are important when considering tradeoffs over long time horizons. Geometrically, Figure 5-8 shows the age profile of WTP for a microrisk reduction in the chance of sudden death in the current period, for three different discount rate assumptions. The influence of different discounting assumptions is greatest among respondents who are currently younger than 55 years of age. For older respondents, the age profiles are relatively robust to different discounting assumptions.

5.6.6 Individual-specific discount rates

In a separate survey that is part of our larger study, we asked a different sample of Knowledge Networks respondents to consider choices among public health risk reduction programs, rather than the private health-risk reduction programs described in this document. As part of that survey, respondents were asked to consider how they might prefer to receive some hypothetical lottery winnings. Bosworth et al. (2011) describes how we develop a model to explain the individual discount rates implied by respondents' answers to whether they would prefer to take a

smaller lump sum payment up front, or wait for the full amount of the lottery winnings to be disbursed in some number of payments over time. The fitted individual discount rate model is:

$$\hat{r}_i = \exp\left(\begin{array}{l} -0.0127*\left(\textit{subjective life expectancy}\right)_i \\ -\ 0.0381*\left(8.81\right)\ -\ 4.49*\left(\textit{age}_i/100\right) \\ +\ 0.432*\left(\textit{age}_i/100\right)*\left(\textit{years of education}\right)_i \\ +\ 0.0157*\textit{female}_i\ +0.522*\textit{female}_i*\textit{nonwhite}_i \\ -0.185*\left(\textit{years of education}\right)_i \\ +\ 0.272\ *\ \left(\textit{income}_i < \$27{,}500\right) \\ -0.396*\left(\textit{Hispanic}_i\right)+\ 0.675 \end{array}\right) \tag{0.37}$$

.

The "subjective recovery likelihood" index variable, not collected in the survey used for this study, is set equal to the sample mean (8.81) in Bosworth et al. (2011) study, since both samples are drawn from the same population. Across the 2407 respondents used in this analysis, the mean calculated individual discount rate is 0.0839. The standard deviation is 0.0306. The minimum and maximum calculated individual discount rates are 0.0300 and 0.4817.

Based on this model of individual discount rate, fitted for that other sample, we calculate point estimates of individual discount rates for this sample, based on the same set of explanatory variables. For this lottery-winnings disbursement choice model, however, the fitted individual-specific discount rates seem somewhat high, averaging about 8.4%. Figure 5-13 shows a histogram describing the distribution of individual discount rates for the 2407 respondents in this study, using the estimated parameters from the model that has been estimated for the public-choices survey respondents.

Table 5-11 compares the results (for the identical specification) for the 5% discounting assumption versus the case with calculated individual discount rates. For ease of comparison, we use the identical set of interaction terms to correct for scenario adjustment/rejection. As expected, the estimated marginal utility-related parameters are systematically different when different discount rates are used to calculate the discounted future years in each health state. The larger (average) discount rates in the model based on individual discount rates will cause the present discounted future life-years in each health state to be smaller, so that to explain the identical choices, we would expect the coefficients to be somewhat larger (at least in the linear and additively separable case; the size difference is somewhat more ambiguous due to the nonlinearities and interaction terms in the model).

Table 5-12 contains the key WTP information. In Column (1) of that table, we show the full set of WTP simulations using the same individual discount rates used in the estimation process for the individual discount rate model. In Columns (2), (3), and (4), however, we over-ride this financially based individual discount rate. We keep the marginal utility-type parameters estimated on the basis of the individual-specific discount rates, but we calculate WTP for each draw from that set of jointly normally distributed parameters by counterfactually simulating a common discount rate for all respondents. These three different sets of WTP amounts are implied by our individual-discount-rate model if we force everyone, instead, to use a 3%, a 5%, or a 7% discount rate for the calculation of social benefits in the absence of the capital market

constraints or other considerations that produce the somewhat larger discount rates elicited through the lottery-winnings disbursement choice question.

The final column on Table 5-12 reproduces the results of the full set of illness profile WTP simulations based on the main model in the paper, which assumes a common 5% discount rate for all respondents even during the estimation phase of the model. *The important finding is that if we wish to base policy choices on the WTP amounts implied by a 5% discount rate, it seems to make very little difference whether we assume this discount rate for everyone, across the board, in the estimation phase, or whether we allow each individual to have a different individual discount rate, then counterfactually simulate a 5% discount rate ex post.*

In the review process for some of the other papers from our study, based on our other Knowledge Networks samples, some of our referees have complained that we have no evidence to confirm that the rates at which consumers discount future health is the same as the rate at which they discount future money. Given this skepticism about the use of the separate discount rates estimated from tradeoffs involving money payments over time, we have opted to feature models with the 5% fixed discount rate assumptions that correspond to the types of discounting assumptions more typically made in the environmental policy arena. Table 5-12 demonstrates that our findings concerning WTP amounts seems to be robust across either this fixed discounting assumption or the transfer of our individual discount rate model from another sample of Knowledge Networks respondents. Thus we are confident that if a 5% discount rate is to be used in other parts of a benefit-cost analysis, and if consistency in discounting assumptions is desired, our 5% discounting results are likely to be appropriate.

For completeness, note that we considered treating the *difference* between the individual-fitted discount rate based on the Bosworth et al. (2011) model as another sort of a scenario adjustment/rejection control variable in our models in the main paper. However, the fitted individual discount rate depends on age and age-squared, as do several of the regressors in our choice specification in this model. Including "correction" variables based on these fitted individual discount rates soaked up much of the explanatory power of the basic utility parameters that involve interaction terms in age and age-squared. Therefore, that approach seems to be inappropriate.

Furthermore, our fitted individual-specific discount rates introduce a lot of sociodemographic heterogeneity in preferences, but ONLY via the exponential discounting parameter. If we are going to allow sociodemographic heterogeneity, there is an argument for allowing these variables to shift all of the utility parameters in the model. While sociodemographic heterogeneity is explored in research we currently have in progress, it would obscure the main points of our basic model to attempt to incorporate all of the potential heterogeneity at once. Importantly, *the attributes of the program choice options are randomly assigned*, other than their dependence upon age (and the dependence of the set of illnesses on gender), so we do not need to worry too much about unobserved heterogeneity producing bias in the marginal utility parameter estimates in our main model.

5.6.7 Age profiles

The main paper includes figures which display the age profiles of mean simulated WTP amounts, along with the 5^{th} and 95^{th} percentiles of the distribution of fitted WTP amounts at each age. For some age levels, some or all of these quantities drop below zero. This can happen because we do not constrain fitted WTP to be non-negative. However, none of our choice

scenarios give the respondent the option to be paid to accept a risk reduction. The most that can be done if a risk-reduction option has no value is for the subject to choose the other alternative or the "no program" option. Thus there is an argument for taking a Tobit-model sort of a perspective and to treat negative fitted values as zero.

In using the Tobit-like interpretation, it would be appropriate to recalculate the mean fitted WTP substituting zeros for any negative simulated values. This would tend to increase the mean WTP to an extent that will reflect the dispersion in the estimates. While it is possible to argue that this strategy is appropriate, we have elected to retain the mean value including negative estimates. When this mean is less than zero, in plotting the age profiles, we set it equal to zero, and we do likewise for the 5^{th} and 95^{th} percentiles. In the main paper, we display Figure 1 (Sudden death now), Figure 2 (1 year sick now, recovery/remission, life-span not affected), Figure 3 (10 years latency, 5 years sick, then die) and Figure 4 (End-of-life effects; half year sick, die half-year early). Figure 5-9 in this Handbook reproduces the age profile for WTP for this last illness profile.

It is worth emphasizing, at this point, that *several* of our main utility parameters are specified as different quadratic (or at least linear) functions of the respondent's current age. The resulting overall age profile for WTP to reduce the risk of a specified illness profile therefore reflects the *combined* effects of all of these age-dependent utility parameters. In some previous studies which have considered the dependence of WTP for risk reductions as a function of age, WTP itself has been allowed to depend directly on age in a more-or-less reduced-form quadratic fashion. Our models are not simply ad hoc specifications where *WTP* is allowed to depend directly on age. The age effects in our model are mediated by our structural model of preferences, and we have revealed that *several* marginal utilities related to future health states depend upon the respondent's current age, and in different ways.

5.6.8 Including an alternative-specific dummy for "either program"

There is no natural ordering to the to risk reduction programs offered in each choice set, since ten illness labels are randomly selected from a possible eleven illnesses or injuries for each gender. These ten illnesses are randomly paired into our three-alternative choices sets (in addition to the "Neither Program" alternative). The order in which the ten illnesses appear for any individual is thus random. Consequently, there is no real argument for alternative-specific dummy variables on the "left" and "right" alternatives in the substantive pair.

However, researchers are sometimes interested in knowing whether there is some unobserved bias either for or against both of the substantive alternatives (versus the status quo "Neither Program" alternative). Testing for such an effect can be done either with a "status quo" dummy variable, or a common "Either Program" dummy associated with both of the offered programs. We use the former option in a model presented in the second column of results in Table 5-7. The first column shows estimates of the same parameters when no status-quo indicator variable is employed in the model.

We find that allowing for there to be some unspecified difference in utility associated with the "Neither Program" options (or, equivalently, an unspecified but opposite difference in utility associated with either of the two risk reduction programs) produces a strongly significant positive point estimate on the additional dummy variable, as well as a large increase in the maximized value of the log-likelihood. We infer that respondents are somewhat inclined to choose one of the two risk reduction programs regardless of the costs and benefits of either program.

A variable such as this is typically employed to capture the *net* effects of phenomena such as "yea-saying" (which would tend to produce a positive coefficient on this variable) or "payment vehicle rejection" (which would tend to produce a negative coefficient). In our data, therefore, it seems that there is some autonomous utility derived from *either* risk-reduction alternative, but not from the status quo. This could be what stated-preference researchers sometimes call "warm glow." It could be that the *benefits* of the risk-reduction programs are perceived to be greater than we describe by some amount unrelated to the specific quantitative attributes used in our choice model. Or perhaps the program *costs* are perceived to be lower than the scenarios state. Unfortunately, as is always the case with these status-quo-effect variables, it is impossible to know exactly what this type of variable is capturing.

We can certainly speculate upon why our model with a status quo effect indicates a systematic preference for either of the program alternatives above and beyond what can be attributed to the time profiles of illness associated with each health risk. In our main paper, we do not control for the illness names associated with each program, relying on the near-independence of the illness names from the illness profiles with which they are associated to preclude any omitted variables bias in the other parameters. We do have another paper focuses specifically on these illness labels and finds that some of them have a statistically significant effect on baseline utility and on the marginal (dis)utilities of sick-time and lost life-years, at least in some cases. These differences result in different WTP for risk reductions as a function of the names of the illnesses. To the extent that there is more about all of the illnesses or survey covers, embodied simply in their names, in addition to the disutility from future time periods in adverse health states, we may miss some of the benefits of risk reductions if we net out this status quo effect.

Other illness attributes which did not bear robust and stable coefficients in any of our empirical models include the information on the mix of moderate and severe pain and disability over the sick-years in the profile, and information about whether hospitalization or minor or major surgery would be required. These variables were randomized, so the fact that their individual coefficients were statistically insignificant in models that do not control for illness names is not a consequence of a high degree of multicollinearity. Of course, the periods of moderate and severe pain, by construction, had to exhaust the specified number of sick-years. More general models may yet pick up statistically significant effects. Without separate controls for these apparently less-important features of each illness profile, the influence of these attributes may also show up as systematic bias against the status quo.

The logic for netting out the autonomous portion of *WTP* captured by the "either program" dummy variable in this study, however, is not entirely clear. When the estimated coefficient is negative, in cases where there are clear reasons to suspect that there is a significant problem with payment vehicle rejection (as in the use of taxes to pay for public goods), it may be defensible to net out the autonomous negative component in *WTP*. In this case, however, respondents appear to be willing to pay some amount for *any* type of health risk reduction program, regardless of the size of its effect or the type of the risk.

Perhaps it is not unreasonable that people *appear* to be willing to pay positive amounts even as the size of the risk change approaches zero. None of our costly programs yields a *zero* risk reduction, so we cannot test for a positive *WTP* even when benefits are zero. Perhaps *WTP* is not exactly proportional to the size of the risk reduction, and this is what the "either program" dummy variable is picking up. For example, perhaps WTP as a function of the size of the risk has a positive intercept, so a substantial component of WTP is induced by the size of the risk

reduction, but a non-zero component is always present, even when the risk reduction goes to zero. Alternatively, WTP may follow a roughly linear trajectory towards a positive intercept as the size of the risk reduction shrinks, until the risk reduction becomes arbitrarily small, at which point WTP jumps discontinuously to zero. Since we have no infinitesimally small risk reductions in our data, we cannot preclude this sort of a trajectory.

We are not entirely convinced, therefore, that it is appropriate to net out the (non)status quo effect in our study, but our featured estimates do so because the reviewers of our paper recommend this strategy. For the intermediate specification employed with an estimating sample that involves additional exclusion restrictions, the results are shown in Table 5-7. In this case, the point estimate of the status quo effect is statistically significantly different from zero. The consequences of including this status quo indicator, and then simulating it to have a value of zero during our *WTP* calculations, are shown in the second column of results in Table 5-8 for this intermediate model. As expected, canceling out this positive lump of utility shared by all risk-reduction programs leads to a modest decrease in *WTP* for each of our basic set of five illness profiles. The first column shows the results for the model without status quo effects. For example, *WTP* to avoid sudden death in the current period drops from $7.59 to $5.95.

However, for the more-general model featured in the main paper (with additional scenario adjustment/rejection control variables and fewer exclusion restrictions), the coefficient on the status quo indicator is no longer statistically significantly different from zero. Nevertheless, to reassure our referees that no distortionary status quo effect is being suppressed, we continue to include this indicator variable (even though statistically insignificant coefficients tend to inflate the confidence intervals for our WTP estimates).

5.6.9 WTP as a function of income levels

Our basic simulations of WTP for a microrisk reduction in the chance of sudden death in the current period are calculated for an income level of $42,000 (2003 U.S. dollars). Variations in WTP as a function of income have been considered one of the few adjustments that may be politically easy to make, since the "value" of risk reductions can be expected to grow over time as incomes grow. Figure shows how this standardized WTP for a microrisk reduction varies systematically with income.

The dependency of WTP on income is fundamental in our model, as is clear from equation (0.21) and (0.22) in the linear-in-income special case, and as is implied by equation (0.23) for the more-general Box-Cox case. In the comprehensive comparisons of implied WTP distributions that accompany each comparison of parameter estimates in the tables at the end of this section, we include at the end of the inventory a set of simulations that shows how higher incomes produce considerably great WTP for health risk reductions.

5.6.10 If respondents expect half as much, or zero, income when sick

In the main model reported in the paper, we assume that $\gamma_1 = 1$ and $\gamma_2 = \gamma_3 = \gamma_4 = 0$. This means we make the assumption that respondents *do not* anticipate having a substantially reduced income, should they suffer the illness or injury described in each illness profile, but they assume that they will not have to pay for the risk-reduction program (diagnostic tests) if they actually get sick from the disease in question. Furthermore, should they die from this illness, they expect to

earn zero income (i.e. to consume no other goods and services) and to be freed from any obligation to pay the cost of the risk reducing program.

However, referees have asked about the effect of different assumptions about "earnings while sick" on our estimated parameters and the implied WTP amounts. It is straightforward to adjust the calculation of the variables for use in the main model to accommodate other assumptions about income levels during illness years. For example, if $\gamma_1 = 1$ and all of the other γ parameters are zero, the expression in (0.6) becomes:

$$\begin{aligned} & f(Y-c)\left[\left(1-\Pi^{AS}\right) pdvc + \Pi^{AS}\left(pdve + pdvr\right)\right] \\ & + f(Y)\left[\Delta\Pi^{AS}\, pdvi - \left\{\left(1-\Pi^{NS}\right) pdvc + \Pi^{NS}\left(pdve + pdvr\right)\right\}\right] \\ & + f(0)\left[\Delta\Pi^{AS}\, pdvl\right] \end{aligned} \tag{0.38}$$

On the other hand, if $\gamma_1 = 0$ the same expression instead simplifies to:

$$\begin{aligned} & f(Y-c)\left[\left(1-\Pi^{AS}\right) pdvc + \Pi^{AS}\left(pdve + pdvr\right)\right] \\ & + f(Y)\left[(-1)\left\{\left(1-\Pi^{NS}\right) pdvc + \Pi^{NS}\left(pdve + pdvr\right)\right\}\right] \\ & + f(0)\left[\Delta\Pi^{AS}\left(pdvi + pdvl\right)\right] \end{aligned} \tag{0.39}$$

Keep in mind that in the linear case, $f(0) = 0$, but in the Box-Cox case, $f(0) = -1/\lambda$.

Going from (0.38) to (0.39) to appreciate what changes when γ_1 changes from 1 to 0, note that the term in square brackets in the first line, which is abbreviated as *cterm* and which appears in the denominator of the WTP formula, is unchanged. The second line loses a term equal to $f(Y)\Delta\Pi^{AS}\, pdvi$ and the third line gains a term equal to $f(0)\Delta\Pi^{AS}\, pdvi$. Thus the change in the numerator term of WTP will be $\left[f(0) - f(Y)\right]\Delta\Pi^{AS}\, pdvi$. This term does not involve program cost. The change in the assumption about income earned while sick results in a change in the *bXterm* variable used in the estimating equation, shifting overall net utility downward (since $f(Y) > f(0)$) but by a small amount because this difference is multiplied by a small negative risk change $\Delta\Pi^{AS}$.

Across individuals, the size of this adjustment term will depend on income level and on the number and future timing of sick-years in the illness profile in question. However, the effect will be to add a positive multiple of *pdvi* to the numerator of the WTP function. If the *bXterm* variable is induced in this way to depend a bit more on the value of *pdvi*, then the corresponding term or terms in the *αterm* expression that involve the variable $\Delta\Pi^{AS}\log\left(pdvi + 1\right)$ will have to share a little more of their explanatory power with *bXterm*. Thus if the model were linear and additively separable in the *pdvi* term, we would expect its coefficient to decrease in absolute value. In Table 5-15, this appears to be the case.

However, we must remember that people's choices don't change. The same behavior merely has to be explained under these different assumptions about expected income while sick. Comparing the parameter estimates across assumptions in Table 5-15, we see that the estimated

slope coefficients adjust to accommodate the different variables. The slope coefficient on the net income variable changes in its third significant figure, but the coefficient on the discounted sick-years term (*pdvi*) becomes more noticeably negative to take up the slack. The change in the estimated parameters absorb the effect of the change in the net income variable, and the overall effect on the numerator in the WTP formula is essentially "a wash." This can be seen in the set of WTP estimates provided in Table 5-16.

Specifically, suppose the respondent expects to earn only $\gamma_1 Y$ during any years when he or she is suffering from a major illness, where $0 \le \gamma_1 \le 1$. Then the formula in expression (0.6) must be adapted. If we retain the assumption that $\gamma_2 = \gamma_3 = \gamma_4 = 0$, this expression becomes:

$$\begin{aligned} & f\left(Y-c\right)\left[\left(1-\Pi^{AS}\right) pdvc + \Pi^{AS}\, pdvp\right] \\ & + f\left(Y\right)\left[\left(-1\right)\left(\left(1-\Pi^{NS}\right) pdvc + \Pi^{NS}\, pdvp\right)\right] + f\left(\gamma_1 Y\right)\left[\Delta\Pi^{AS}\, pdvi\right] \end{aligned} \tag{0.40}$$

For $\gamma_1 = 0.5$ and $\gamma_1 = 0$, the second and third models in Table 5-15, and the second and third sets of WTP simulations in Table 5-16 provide the details concerning the effects of these adjustments on our estimates. The impact of these changes on WTP is extremely small.

5.6.11 If respondents perceive other costs in addition to those quoted

One reviewer of the main paper was concerned that some respondents may have treated the stated costs of each program as less than the full opportunity cost that would be involved if they chose to participate. On **Form 17**, we state specifically that the risk reduction programs in question would *not* involve "uncomfortable procedures." We do state that "Your participation in a program would cost you money." These programs would not be covered by the respondent's current health insurance. "These higher costs might take the form of a co-payment when you visit your doctor or higher monthly health insurance costs." "To make it easier to compare, we present all costs as monthly costs, and also as annual costs. You would need to pay for, and participate in, a program for the next __ years to get its benefits." (The precise number of years corresponded to that individual's current age and nominal gender-specific life expectancy.) We did not explicitly limit the cost of the *program* to simply the cost of the *test*. Instead, we were careful to refer to the "cost of the program" (where the programs are described on **Form 17** as involving prescribed "medication and life-style changes that reduce your risk of getting the illness").

Earlier in the survey, however, on **Form 7**, we specifically asked respondents to consider the difficulty of making life-style changes. We asked them: "Changing your lifestyle or habits can be difficult because it requires time, money, and effort. How difficult would it be for you to do the following things?" The listed options included the following measures: drink less alcohol, quit smoking, eat a healthier diet, see a doctor more regularly, exercise more, lose weight, use a seatbelt more. We went through one phase of survey development with language in the instrument where we tried to explain the idea of the monetized disutility of the tests themselves, and opportunity costs and the full cost of time. However, without getting into discussions of the value of travel time to the doctor's office and the pharmacy, and the prospective disutility of a new exercise regimen or dietary restrictions, there seemed to be no happy medium, so we opted

for a minimalist approach. Perhaps there would have been a better option, but we could not see it at the time. To meet the length/duration restrictions under our contract with Knowledge Networks, of course, it was necessary to prune many things out of the survey that we were keenly interested to include. This is a frequent problem with survey research in general. One's claim on the respondent's time is a finite resource.

In response to this concern, however, we have investigated additional models where we allow the estimated marginal utility of net income to depend on the respondent's answers to our questions about the difficulty of accomplishing lifestyle changes. We take advantage of the wording on **Form 7** in the question: "Changing your lifestyle or habits can be difficult because it requires time, money, and effort. How difficult would it be for you to do the following things?" A slight complication is that respondents were only asked about each of these things if they responded on **Form 6** that there was still at least some room for them to reduce their health risks by improving their lifestyle or habits in that particular way. We assume that if the individual reports no room to improve along any particular dimension, then it would be very hard at the margin for them to improve any further on this dimension. (Cleaning up a few of your bad habits may be relatively easy, but getting rid of all of them might be tough.)

However, if there is still room to improve on one or more dimensions, and respondents report that it would be easy or difficult for them to do so, this is the notion we wish to capture. We construct a crude variable to measure "ease of improving health habits." For each type of the seven health habits identified on **Form 6** and **Form 7**, we build two variables. One is prefixed by "improve_" and measures "opportunity for improvement" with ratings that vary from 0 = "no opportunity for improvement" to 4 = "much room to improve." The second variable is prefixed by "easy_" and measures the ease with which these available improvements in health habits could be accomplished. For this variable, we have inverted the question about how *difficult* it would be to make improvements. For our "easy_" variables, the ratings are coded as 0 = "hard to improve" to 4 = "easy to improve."

For each of the seven health habits, we construct an interaction between the "improve_" and "easy_" variables. This interaction term is zero if the individual has no opportunity to improve *or* if they do, but it would be very hard for them to do so. This interaction term takes on a larger value (to a maximum value of 16) if there is lots of room for the individual to improve their health habits and they believe it would be easy to do so. Acknowledging the degree of approximation involved in the use of ratings, and the different metrics across the different questions, we then forge ahead and add these interacted ratings across all seven types of health habits to generate a variable that may serve as a proxy for the likely psychic or non-pecuniary costs to the individual if they need to make "lifestyle changes" in addition to paying for the annual pin-prick blood test in the choice scenarios.

The maximum value for our constructed indicator is 16x7 = 112. It measures "ease of making lifestyle changes." We desire a variable that will be larger if the implicit costs to the individual of making these changes is larger, so we subtract our indicator from 112 to convert it into an indicator called $hard_i$ which proxies for the "difficulty of making lifestyle changes." As a further complication, however, not all respondents answered all of the questions on **Form 6** and **Form 7**, so we create an indicator for whether information was missing. 1,724 of our 1,801 respondents (in our sample based on three exclusion criteria) provided sufficient information to build this variable. We thus use a second indicator variable to control for data availability.

Now we simplify the intuition by supposing that the indirect utility difference that drives program choices is linear in net income and we don't need to worry about the pattern of net

income across the uncertain prospects of getting sick or remaining healthy. In that simple case, $\beta(Y-c)-\beta(Y)=\beta(-c)$. Suppose costs are perceived as systematically higher than what is stated in the choice scenario, say $c\theta$, where $\theta>1$. If respondents are reacting to this larger cost, but we control only for c, then we will actually be estimating $(\beta\theta)(-c)$, and the apparent "marginal utility of net income" coefficient will be too large. This coefficient forms the denominator of the WTP function, so a too-large value will lead to a WTP estimate that is too small. People who look like they are unwilling to pay the amount stated in the choice scenario are actually unwilling to pay the larger implicit cost, rather than the actual stated amount mentioned in the choice scenario. Failure to accommodate these other implicit costs will lead to underestimates of WTP.

We incorporate our new variable, $hard_i$, along with the indicator for its availability, into our model by allowing these two variables to shift the β coefficient. The slope coefficient on the interaction with the indicator variable is insignificant, but the slope coefficient on the interaction with $hard_i$ is positive and strongly significant. If we estimate β as a scalar, its point estimate is 0.0139. (We do not constrain the systematically varying version of this parameter to be positive, so a few negative values result.) Figure 5-11 shows the range of implied values for the β parameter in this more-general model. The mean of these fitted values is 0.0145 and the median is 0.0154.

Thus there exists a range of perceived difficulties of making life-style changes among our respondents. The values of the $hard_i$ variable range from 0 through 112 , with a median of 92 and an interquartile range of 82 through 99. For people who perceive life-style changes as relatively more difficult (i.e. those who may consider other implicit costs associated with each risk reduction program), the marginal utility of income is estimated to be higher, which would imply a lower WTP for the risk reduction programs in the choice scenarios. For people who perceive life-style changes as relatively easier, the marginal utility of income is estimated to be lower, which would imply a higher WTP for the risk-reduction programs in the choice scenarios.

As an alternative, we could build the $hard_i$ variable using *only* the information on how easy it would be to improve each health-related behavior on the list (i.e. without the information on whether the individual has room to improve). When we do this, the implied β parameters display the range shown in Figure 5-12. In this case, there are fewer negative fitted values, but the results are qualitatively the same. The slope coefficient on the interaction between the $hard_i$ variable and the net income term is positive and strongly statistically significant.

The relevant question, now, is "what would people have been willing to pay had they believed that the quoted cost on the survey was the full cost of the program—i.e. that there were no additional costs associated with the difficulty of complying with the lifestyle changes that might be required?" It might be tempting to simulate the value of the marginal utility of income parameter for the case where everyone believes that it is trivially easy to implement life-style changes. This would correspond to the counterfactual where nobody perceives any implicit costs of this variety in addition to the cost of having the test.

We had intended to do this sort of thing in our analysis, which was why we collected the information on **Form 6** and **Form 7**. However, we did not anticipate that respondents might view "lifestyle changes" in two separate ways. We expected that people would view them as necessary *complements* to the health testing programs described in the choice scenarios. This is

the implicit assumption behind the concern that respondents will impute other costs to each program besides just the cost stated in the survey question. However, it may actually be the case that respondents view the testing programs in the survey as *substitutes* for the lifestyle changes that they know they should really be trying to make. If they perceive that participation in these testing programs will allow them the luxury to continue with their current poor health habits but still lower their health risks, they may actually express greater demand—because the perceived benefits are greater than just the reduction of health risks.

This makes things considerably more complicated. If we were to simulate a situation where everyone found it perfectly easy to implement any required life-style changes that would be required along with the testing program, the marginal utility parameter for income would be vastly smaller, causing the inferred WTP for these programs to be vastly bigger. But here's the catch: if lifestyle changes were easy, the "price of a substitute" for the testing program would also be dramatically smaller, which would *decrease the demand* for the testing programs. People could simply change their health habits and they would have no need for the testing program. Thus it seems highly inappropriate to consider any adjustments to the stated cost of the program without making corresponding adjustment to the price of substitutes. Clearly, more research is needed, and it should focus on this "complements versus substitutes" distinction.

Incidentally, we do have some evidence, in other work with these data, for the "substitutes" possibility. In our research concerning the disease labels, non-smokers are willing to pay very little for tests to reduce their risk of lung cancer or respiratory disease, whereas smokers are willing to pay amounts for these two illnesses that substantially exceed the WTP amounts measured for all other illnesses for the general population. In this case, the substitution effect appears to dominate very strongly.

5.6.12 Effects of risk aversion in preferences (i.e. curvature in $f(\cdot)$)

In a linear model such as that described in Section 5.1.5.1, the baseline level of the health risk drops out of the expression for WTP. With respect to the risks involved, then, WTP thus depends only upon the size of the risk *reduction*, $\Delta\Pi_i^{AS}$ (and it is strictly proportional to the size of the risk reduction, as in equation (0.19) if we disregard the error term).

If the model is non-linear in net income, however, the baseline level of the risk does not drop out of expression for WTP except in special circumstances. In the general case, due to the presence of $\lambda \neq 0$ in equation (0.23), the *yterm* expressions, as defined in equation (0.7), will involve Π_i^{NS} as well as $\Delta\Pi_i^{AS}$ (where $\Delta\Pi_i^{AS} = \Pi_i^{AS} - \Pi_i^{NS}$ so at most two of these three terms are independent). Thus the calculated "average WTP for a microrisk reduction, calculated for a given sized risk reduction based on an initial baseline risk" will depend on that baseline risk if there is any curvature in the utility function in the direction of net income. Only in the case of no curvature is the average WTP for a microrisk reduction identical regardless of the starting point of this risk change.

When we allow for curvature to the extent of $\lambda = 0.45$, it is therefore important to consider how the implied average WTP for a microrisk reduction might depend upon the starting point and the size of the risk reduction. Thus we need to use our estimated preference parameters for the Box-Cox specification in our main paper to simulate the value of WTP for a microrisk reduction as a function of different baseline levels of risk (and for good measure, for different-sized risk reductions from those baselines).

5.6.12.1 Special case: Risk reduction provided as a public good

Credible risk reduction scenarios to elicit private WTP for privately provided risk reductions motivated our characterization of the risk reductions in our choice scenarios as diagnostic medical tests (in the case of illnesses) or retrofitted equipment in private automobiles (in the case of traffic accidents). On **Form 21** of the survey, we were careful to tell respondents that

> "You would need to pay for, and participate in, a program for the next <<remaining life>> years to get its benefits."

In the case of the diagnostic tests, however, it was likely implausible to respondents that they would have to continue paying for a diagnostic test to reduce their risk of getting a particular illness if they are unambiguously suffering from that illness at some future point in time. Due to these considerations, we set $\gamma_3 = 0$. In the case of traffic accidents, however, the logical assumption may be less clear. However, if the vehicle in question was "totaled" in the accident, the cost of the risk-reducing upgrade, if it was chosen, would be paid off by insurance along with the rest of the vehicle. For the period of "moderate" or "severe" pain and disability associated with the accident, perhaps no replacement vehicle has yet been purchased. Based on the assumption of a significant injury accident, we impose the $\gamma_3 = 0$ assumption during estimation for traffic accident risks as well.

However, we wish to use our estimated WTP amounts as measures of the demand for publicly provided risk reductions. In these cases, suppose that the risk reductions are funded by taxes or by regulations that result in higher production costs for consumer goods, or lower wages, or lower investment returns. In cases such as these, the obligation to pay for the risk reductions would not go away if the individual in questions actually suffered the illness or injury. In simulating WTP amounts, therefore, it may be appropriate to force the assumption that $\gamma_3 = 1$. When this assumption is imposed during the simulation of WTP, we will refer to that WTP estimate as WTP for a "public" risk reduction. If we impose the assumption of $\gamma_3 = 0$ during the simulation of WTP, we will refer to this WTP estimate as WTP for a "private" risk reduction.

Table 5-17 employs the utility parameter estimates from our main model (where indirect utility is Box-Cox in net income with a parameter of 0.45) and shows simulated WTP distributions for private and public risk reductions under a variety of alternative baseline risks and risk reductions. The differences in the WTP estimates across these alternative assumptions are very minimal. Table 5-17 shows the (expected) results from a similar exercise where the utility function is specified as linear in net income. In this case, we expect no differences in the WTP estimates for a microrisk reduction. Aside from what we believe are minor rounding errors in our algorithm, on the order of a couple of pennies at most, this appears to be true.

It may be especially reassuring to compare the analogous risk reductions for our Box-Cox specification in Table 5-17 and a linear approximation (not reported here). These differences can only be characterized as "very minor" relative to the 90% interval for each estimate. So despite the improvement in the log-likelihood values due to the introduction of the Box-Cox parameter, the degree of curvature in the preference function over the relevant range is small and appears to be of relatively little consequence, at least for these particular simulations.

5.6.12.2 Special case: Simulated WTP with constant net income for nominal lifetime

Even if utility is diminishing in net income, as it is in the Box-Cox specification with $\lambda = 0.45$, there is one case where simulated WTP loses its dependence upon the baseline level of risk. We can employ our standard assumptions about income and costs while sick or dead in the estimation phase. However, if respondents are believed to *ignore the possibility of being relieved of their responsibility to pay the costs of the program when they are sick or dead,* and if they *fail to think about income losses when they are sick or dead,* we can simulate WTP under conditions where $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 1$. In this case, the key terms in equation (0.7) become:

$$\begin{aligned} cterm &= pdvc \\ yterm1 &= (-1)\left\{\left(1-\Pi^{NS}\right) pdvc + \Pi^{NS}\left(pdve + pdvr\right)\right\} \\ yterm2 &= \left(-\Pi^{NS}\right) pdvi \\ yterm3 &= \left(-\Pi^{NS}\right) pdvl \end{aligned} \tag{0.41}$$

So that equation (0.23) can be simplified as follows:

$$c_i^j * = Y - \left(1 + \lambda\left\{\beta\left(\frac{Y_i^{\lambda} - 1}{\lambda}\right) - \frac{\alpha term_i^j}{pdvc_i}\Delta\Pi^{jS} - \frac{\varepsilon_i^j}{pdvc_i}\right\}\right)^{1/\lambda} \tag{0.42}$$

since it is possible under these conditions to aggregate or simplify so many of the terms in the WTP expression. Only the $\Delta\Pi_i^{AS}$ probability term now remains, since the various terms in the absolute levels of any of the probabilities drop out. Willingness to pay is thus no longer dependent upon baseline risk levels, which may be convenient. In situations where the consumer sees the cost of the program and their own income remaining constant for the duration of their remaining nominal life expectancy, the formula in equation (0.42) may be the appropriate basis for WTP calculations. We call this the "flat lifetime net income" assumption.

5.6.13 Effect of position in choice order

Referees have raised the question of order effects across the five choice sets presented to each respondent. Fortunately, the order of the illness names was randomized for each respondent, as were the attributes of each illness profile (subject to minor plausibility constraints and the constraints imposed by the respondent's gender and age). Each set of choice scenarios was essentially unique. Thus there can be no systematic effects of the order of the named illnesses or the patterns in their illness profiles. Our estimated marginal utility parameters are essentially the average effects of attribute levels across the range of illnesses and the range of illness profiles used in this survey.

Still, one might be interested in the effects of "choice number" on the implied preference parameters. Are the preferences implied by the first choice set respondents saw systematically different from the preferences implied by later choice sets presented to each respondent. Any such effects could reflect the net effect of learning, any evolving choice heuristics, or fatigue.

Table 5-18 shows both the parameter estimates and then the simulated WTP estimates for a sequence of six models. For Models 1 through 5, we introduce two new classes of control variables that we interact with the basic variables in our specification. The first set of controls involves statistically significant interactions of these variables with the deviation of time-on-choice from the overall average choice duration (for a distribution trimmed of its extreme outliers), *timedev*, and *timedev* squared. The case of interest is the one where all time-on-choice deviations are zero. The second set of controls involves statistically significant interactions between the basic variables and linear and quadratic terms in the choice number, measured as a deviation from the "base choice," which can be set to be any of the five choice sets. Where dictated by the data, this form allows for a U-shaped (or inverse U-shaped) profile for each parameter in the main part of the model, as the choice number changes.

Model 0, displayed first, differs from the others in that it *excludes* a set of controls for the deviation of time-on-choice from the overall average choice duration (for a distribution trimmed of its extreme outliers), *timedev*, and *timedev* squared. Model 0 also excludes indicators for the choice number measured as a deviation from the desired "base" choice, and this deviation squared.

The other Models (1 through 5) normalize the base choice on each different choice number. The reason for this approach, as with the suite of scenario adjustment/rejection controls used in this study, is to permit us to assign zero values to the case of interest, so that all incidental controls drop out of the model and we are left with the basic specification as the utility-difference equation that should prevail under the desired conditions.

Among the simulated WTP values at the bottom of the table, then, the first column shows WTP amounts with no controls for time-on-task or for choice number and the remaining five columns show WTP estimates normalized on average choice durations and on the choice number in question. Differences across columns are sometimes discernible, but for the most part, the numbers are fairly consistent.

In the lower portion of Table 5-18, it can be seen that the implied WTP amounts for the first benchmark illness profile, "sudden death now," tend to decline as we progress from the first choice as the baseline to the fifth choice as the baseline, although all of the simulated confidence intervals have a substantial degree of overlap. In contrast, for the two benchmark illness profiles with only sick-time and no lost life-years, the implied WTP amounts increase from the first choice to the fifth choice, although the confidence intervals again overlap substantially. This may suggest that as the respondent proceeds through the choice sets, sick-time may become more salient and lost life-years may become less salient. For the two benchmark illness profiles that involve both sick-time and lost life-years, however, the combined effect of these two apparent tendencies leads to different trends in WTP across choice sets, depending upon which trend dominates.

Ultimately, we have decided to rely upon the randomization of illness profiles and illness names across choice sets and to estimate "overall" preference parameters without the *timedev* or *choice number* interaction terms. There would appear to be no a priori reason to normalize choices on the trimmed mean of all observed choice times, or to presume that any particular

choice number guarantees sufficient learning and sufficiently minimal fatigue to make its implied preferences more valid than those calculated for some other baseline choice number.

5.7 Section 5 Tables

Table 5-1 Net income for different health states and program choices
(See Section 5.1.1)

Indirect utility, Probability	Pre-illness/ latency ("e")	Illness/ injury time ("i")	Recovered/ remission ("r")	Lost life-years("l")
$V_t^{AH}, \left(1-\Pi^{AS}\right)$	*Y-c*	*Y-c*	*Y-c*	*Y-c*
V_t^{AS}, Π^{AS}	*Y-c*	$\gamma_1 Y - \gamma_3 c$	*Y-c*	$\gamma_2 Y - \gamma_4 c$
$V_t^{NH}, \left(1-\Pi^{NS}\right)$	*Y*	*Y*	*Y*	*Y*
V_t^{NS}, Π^{NS}	*Y*	$\gamma_1 Y$	*Y*	$\gamma_2 Y$
Discounted time in health state:	*pdve*	*pdvi*	*pdvr*	*pdvl*

The γ parameters reflect the investigator's best assessment of the fractions of income or program costs respondents typically assumed they would receive/pay during any sick-years and after their death. For indirect utility functions which are nonlinear in net income, such as the Box-Cox transformed specification used in the main paper, it is necessary for tractability that the parameters γ_3 and γ_4 take on no values other than 0 or 1. The parameters γ_1 and γ_2, however, may take on any value between 0 to 1 inclusive.

Table 5-2 Utility from *one* period in each health state, by program choice
(See Section 5.1.2)

Indirect utility, Probability	Pre-illness/ latency ("e")	Illness/ injury time ("i")	Recovered/ remission ("r")	Lost life-years("l")
$V_t^{AH}, \left(1-\Pi^{AS}\right)$	α_0	α_0	α_0	α_0
V_t^{AS}, Π^{AS}	α_0	α_1	α_2	α_3
$V_t^{NH}, \left(1-\Pi^{NS}\right)$	α_0	α_0	α_0	α_0
V_t^{NS}, Π^{NS}	α_0	α_1	α_2	α_3
Discounted time in health state:	*pdve*	*pdvi*	*pdvr*	*pdvl*

Table 5-3 Simple model with error dispersion scaled by disease indicators

(See Section 5.4. Simple preliminary specification. Based on data using three main exclusion criteria and no scenario adjustment/rejection controls.)

	Homoscedastic model	Heteroskedastic model	
Variable:	Coef.	Coef.	t-test
Linear term in net income	4.53[a]	5.62[b]	2.22**
Quadratic term in net income	-1.77	-2.17	-2.02**
Sick-years term	-8.81	-9.19	-1.89*
Recovered/remission years	-8.23	-12.35	-1.71*
Lost life-years term	-8.38	-11.37	-2.17**
Error dispersion shifters (relative to heart disease):			
Breast cancer		-0.473	-1.59
Prostate cancer		0.333	0.75
Colon cancer		-0.0486	-0.15
Lung cancer		-0.133	-0.42
Skin cancer		0.341	0.89
Heart attack		0.146	0.44
Stroke		0.0991	0.29
Respiratory disease		0.597	1.37
Traffic accident		0.570	1.39
Diabetes		0.234	0.67
Alzheimer's disease		0.00913	0.03

[a] Average coefficients normalized on average error dispersion across all illness categories.

[b] Average coefficients, normalize instead on the dispersion in the base category, heart disease, in the heteroskedastic model.

Table 5-4 Ad hoc models versus simplest structural model

(See Section 5.6.1. Individuals = 1,801, completed choice sets = 7,520; three main exclusion criteria, no scenario adjustment/rejection controls, no selection correction, fixed effects conditional logit estimates[a])

	Model 1 Ad hoc	Model 2 Ad hoc	Model 3 Structural
Monthly cost of program	-0.007581 (9.63)***[b]	-0.00749 (9.48)***	-
Risk reduction: $\left\|\Delta\Pi_i^A\right\|$	89.27 (9.95)***	57.6 (5.77)***	-
Sick-years	-	0.00880 (3.85)***	-
Lost life-years	-	0.0114 (7.13)***	-
$\left(\left(Y_i - c_i^j\right)^{(.42)} cterm_i^j - \left(Y_i\right)^{(.42)} yterm_i^j\right)$	-	-	0.0144[c] (9.43)***
$\Delta\Pi_i^{jS} pdvi_i^j$ (sick-years term)	-	-	-11.0 (4.90)***
$\Delta\Pi_i^{jS} pdvr_i^j$ (remission-years term)	-	-	-11.8 (2.49)***
$\Delta\Pi_i^{jS} pdvl_i^j$ (lost life-years term)	-	-	-9.92 (5.11)***
Maximized log-likelihood	-11735.125	-11706.105	-11733.32

[a] Each respondent is asked to consider five choice sets, so these are panel data. We use the maximum likelihood estimator that biostatisticians and epidemiologists call "conditional logistic regression for matched case-control groups" and that economists and other social scientists call "fixed-effects logit for panel data." The estimator is coded as "clogit" in the Stata software package. See Greene (2008, p. 800-806).

[b] Absolute asymptotic t-test statistics in parentheses (***=statistically significant at the 1% level; **=statistically significant at the 5% level).

[c] The superscript in parentheses denotes a Box-Cox transformation with the indicated parameter value: $X^{(\lambda)} = \left(X^{\lambda} - 1\right)/\lambda$. The value of 0.42 for λ was determined by a line-search in our initial detailed models. Standard errors are of course conditional on this value for λ. We have previously used square root or quadratic transformations as approximations that dominated either a linear or a logarithmic function for the net income variable, but any of these transformations produces a very strongly statistically significant coefficient.

Table 5-5 Effect of model generalizations on key indirect utility parameters

(See Sections 5.6.2 and 5.6.3. Explorations based on sample with only one exclusion criterion and a subset of the eventual scenario adjustment/rejection controls.)

Basic variables:	1 Linear in logs (4 MU parms)	2 +Higher order log terms	3 +Selectivity correction	4 +Three types of scenario adj.	**5** **+Two types of** **scenario adj.** **Status quo eff.**
$\left(Y_i - c_i^j\right)^{(0.45)} cterm_i^j - \left(Y_i\right)^{(0.45)} yterm_i^j$	0.01044	0.01062	0.01059	0.01458	.01141
	(10.48)***	(9.60)***	(9.58)***	(7.69)***	(9.86)***
$\Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$	-26.73	-46.65	-49.05	-46.66	-59.82
	(4.65)***	(5.37)***	(5.61)***	(3.84)***	(5.31)***
$\Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	-22.61	44.37	46.27	41.25	67.49
	(2.43)**	(1.36)	(1.41)	(1.15)	(2.01)**
… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	-	-1.229	-1.267	-2.701	-2.205
		(1.96)**	(2.02)**	(3.76)***	(3.05)***
$\Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-28.27	-591.5	-589.4	-1240	-549.7
	(5.70)***	(3.30)***	(3.29)***	(5.85)***	(2.99)***
… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-	20.49	20.39	58.17	20.42
		(2.82)***	(2.81)***	(6.84)***	(2.75)***
… $age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-	-0.186	-0.1847	-0.4961	-.1924
		(2.68)***	(2.66)***	(6.21)***	(2.71)***
$\Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	-	206.3	205.1	444.1	176.2
		(2.49)**	(2.48)**	(4.55)***	(2.07)**
… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	-	-7.846	-7.783	-21.56	-7.521
		(2.33)**	(2.31)**	(5.48)***	(2.18)**
… $age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	-	0.07367	0.07297	0.1839	.07554
		(2.26)**	(2.24)**	(4.90)***	(2.25)**
$\Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	-	102.1	99.36	143.3	113.3
		(1.4)	(1.36)	(1.77)*	(1.51)
… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	-	-4.461	-4.367	-7.667	-4.266

		(1.57)	(1.53)	(2.46)**	(1.46)
… $age_{i0}^2 \times \Delta\Pi_i^{jS}\left[\log\left(pdvi_i^j+1\right)\right]\times\left[\log\left(pdvl_i^j+1\right)\right]$	-	0.05604 (2.10)**	0.05528 (2.07)**	0.07168 (2.46)**	.04813 (1.74)*
Status quo effect:					
1(neither program)	-	-	-	-	-.2339 (5.52)***
Systematic selection correction term:					
$\left[P(sel_i)-\overline{P}\right]\times\Delta\Pi_i^{jS}\left[\log\left(pdvi_i^j+1\right)\right]$	-	-	3.440 (2.39)**	3.947 (2.48)**	3.285 (2.22)**
*Scenario adjustment terms (*variable acronyms only*):*					
dilog_agenow_bn	-	-	-	18.63 (4.08)***	22.4 (5.47)***
dilog_agenow2_bn	-	-	-	-0.2199 (3.34)***	-.2709 (4.57)***
dllog_bn	-	-	-	1116 (3.02)***	-
dllog_agenow_bn	-	-	-	-49.8 (2.93)***	-
dllog_agenow2_bn	-	-	-	0.4945 (2.61)***	.1339 (4.03)***
dllog2_bn	-	-	-	-	707.2 (6.46)***
dllog2_agenow_bn	-	-	-	13.96 (2.91)***	-13.22 (5.74)***
dllog2_agenow2_bn	-	-	-	-0.1947 (2.58)***	-
didllog_agenow_bn	-	-	-	-21.37	-26.57

				(4.44)***	(5.27)***
didllog_agenow2_bn	-	-	-	0.3151	.3668
				(4.19)***	(4.70)***
dilog_agenow_bdpos	-	-	-	0.4289	-
				(3.44)***	
dilog_agenow2_bdpos	-	-	-	-0.006616	-
				(3.12)***	
drlog_bdpos	-	-	-	9.646	-
				(2.06)**	
dllog_agenow2_bdpos	-	-	-	0.002042	-
				(3.08)***	
drdllog_bdpos	-	-	-	-35.03	-
				(3.19)***	
drdllog_agenow_bdpos	-	-	-	0.6964	-
				(2.76)***	
b7term_bdneg	-	-	-	0.0005206	-
				(4.59)***	
dilog_bdneg	-	-	-	8.679	-
				(7.96)***	
dllog_agenow2_bdneg	-	-	-	0.007984	-
				(9.15)***	
dllog2_agenow_bdneg	-	-	-	0.1441	-
				(3.36)***	
dllog2_agenow2_bdneg	-	-	-	-0.003228	-
				(3.42)***	
didllog_bdneg	-	-	-	-4.245	-
				(2.97)***	
dllog_ldpos	-	-	-	118.3	-
				(2.31)**	

dllog_agenow_ldpos	-	-	-	-5.338 (2.25)**	-
dllog_agenow2_ldpos	-	-	-	0.05675 (2.16)**	-
dllog2_ldpos	-	-	-	-55.51 (2.18)**	-
dllog2_agenow_ldpos	-	-	-	2.575 (2.19)**	-
dllog2_agenow2_ldpos	-	-	-	-0.0277 (2.12)**	-
b7term_ldneg	-	-	-	-0.0004341 (3.26)***	-
drlog_agenow_ldneg	-	-	-	-0.06801 (2.41)**	-
dilog_agenow2_logldneg	-	-	-	-	.003109 (2.29)**
drlog_agenow_logldneg	-	-	-	-	.3651 (2.21)**
Observations (after three types of exclusions)	22560	22560	22560	22560	**22560**
Log L	-11719.832	-11686.085	-11683.11	-10901.52	-11471.184

Table 5-6 Effect of model generalizations on average WTP for a microrisk reduction
(See Sections 5.6.2 and 5.6.3. Based on the parameter estimates in Table 5-5.)

	1 Linear in logs (4 MU parms)	2 +Higher order log terms	3 +Selectivity correction	4 +Three types of scenario adj.	**5** **+Two types of scenario adj. Status quo eff.**
Income= $42,000					
Now 45: Sudden death at 45	$ 4.41[a] (3.61, 5.27)	$ 5.44 (3.67, 7.33)	$ 5.42 (3.49, 7.42)	$ 7.56 (5.41, 10.02)	$ 5.96 (4.04, 8.11)
Now 45: at 45: 1 yr sick; recov	2.88 (1.44, 4.36)	2.18 (0.69, 3.75)	2.26 (0.78, 3.91)	5.46 (3.63, 7.38)	4.04 (1.94, 6.42)
Now 45: at 45: 5 yrs sick; recov	3.61 (2.2, 5.06)	3.59 (2.04, 5.14)	3.76 (2.28, 5.34)	6.33 (4.52, 8.2)	5.62 (3.46, 7.92)
Now 45: at 45: 1 yr sick; then die	4.87 (4.12, 5.66)	5.37 (3.69, 7.22)	5.4 (3.63, 7.4)	10.24 (7.98, 13.13)	5.91 (4.13, 7.9)
Now 45: at 45: 5 yrs sick; then die	5.23 (4.48, 6.01)	5.11 (3.24, 7.3)	5.23 (3.25, 7.53)	12.03 (9.37, 15.25)	5.54 (3.59, 7.58)
Income= $25,000					
Now 45: Sudden death at 45	3.06 (2.46, 3.7)	3.83 (2.5, 5.25)	3.81 (2.37, 5.32)	5.42 (3.81, 7.26)	4.22 (2.78, 5.83)
Now 45: at 45: 1 yr sick; recov	2.17 (1.09, 3.28)	1.64 (0.52, 2.82)	1.7 (0.59, 2.93)	4.1 (2.73, 5.54)	3.03 (1.46, 4.82)
Now 45: at 45: 5 yrs sick; recov	2.71 (1.66, 3.8)	2.69 (1.53, 3.86)	2.83 (1.71, 4.01)	4.74 (3.39, 6.15)	4.22 (2.6, 5.94)
Now 45: at 45: 1 yr sick; then die	3.42 (2.86, 4.01)	3.79 (2.53, 5.18)	3.82 (2.48, 5.31)	7.43 (5.75, 9.59)	4.2 (2.86, 5.69)
Now 45: at 45: 5 yrs sick; then die	3.74 (3.17, 4.32)	3.65 (2.25, 5.29)	3.73 (2.25, 5.46)	8.82 (6.84, 11.22)	3.97 (2.51, 5.5)
Income= $67,500					
Now 45: Sudden death at 45	6.23 (5.2, 7.35)	7.56 (5.26, 10.03)	7.54 (5.04, 10.15)	10.33 (7.53, 13.54)	8.25 (5.75, 11.04)
Now 45: at 45: 1 yr sick; recov	3.74 (1.87, 5.67)	2.82 (0.89, 4.87)	2.93 (1.01, 5.08)	7.1 (4.71, 9.6)	5.25 (2.51, 8.35)
Now 45: at 45: 5 yrs sick; recov	4.68 (2.86, 6.58)	4.66 (2.65, 6.68)	4.89 (2.96, 6.94)	8.22 (5.87, 10.67)	7.3 (4.49, 10.29)
Now 45: at 45: 1 yr sick; then die	6.8 (5.83, 7.83)	7.45 (5.26, 9.86)	7.49 (5.18, 10.09)	13.8 (10.85, 17.58)	8.15 (5.83, 10.74)
Now 45: at 45: 5 yrs sick; then die	7.18 (6.19, 8.18)	7.01 (4.59, 9.87)	7.17 (4.59, 10.17)	16.04 (12.57, 20.25)	7.57 (5.03, 10.23)
Latency (Income= $42K)					
Now 35: Sudden death now	4.55 (3.74, 5.42)	5.3 (3.06, 7.79)	5.29 (2.84, 7.83)	6.75 (4.32, 9.34)	6.69 (4.28, 9.27)
Now 35: Sudden death at 40	3.92 (3.17, 4.72)	5.1 (3.32, 7.09)	5.1 (3.07, 7.2)	4.5 (2.62, 6.43)	6 (4.06, 8.07)
Now 35: Sudden death at 50	2.88 (2.27, 3.53)	4.65 (3.41, 6)	4.68 (3.38, 6.18)	0.95 (-0.23, 2.14)	4.77 (3.45, 6.25)
Now 35: Sudden death at	2.03	4.04	4.08	-1.35	3.64

60	(1.57, 2.54)	(3.03, 5.13)	(3.04, 5.27)	(-2.5, -0.3)	(2.57, 4.76)
Now 35: Sudden death at 70	1.3 (0.98, 1.64)	3.12 (2.24, 4.04)	3.16 (2.3, 4.14)	-2.37 (-3.57, -1.34)	2.5 (1.61, 3.43)
Now 35: Sudden death at 80	0.59 (0.44, 0.76)	1.71 (1.16, 2.33)	1.74 (1.19, 2.35)	-1.89 (-2.72, -1.18)	1.23 (.66, 1.82)
Now 35: now: 1 yr sick; recov	2.88 (1.41, 4.39)	1.09 (-1.09, 3.17)	1.13 (-1, 3.41)	4.31 (2.36, 6.47)	2.21 (-.07, 4.58)
Now 35: at 40: 1 yr sick; recov	2.6 (1.26, 3.96)	0.94 (-1.06, 2.84)	0.97 (-0.99, 3.04)	3.85 (2.07, 5.82)	1.94 (-.14, 4.08)
Now 35: at 50: 1 yr sick; recov	2.04 (0.97, 3.16)	0.7 (-0.91, 2.23)	0.72 (-0.88, 2.38)	2.95 (1.52, 4.55)	1.46 (-.24, 3.17)
Now 35: at 60: 1 yr sick; recov	1.52 (0.71, 2.35)	0.53 (-0.67, 1.68)	0.54 (-0.64, 1.77)	2.08 (1.01, 3.28)	1.07 (-.21, 2.36)
Now 35: at 70: 1 yr sick; recov	1.01 (0.48, 1.55)	0.43 (-0.34, 1.18)	0.44 (-0.32, 1.23)	1.23 (0.54, 1.98)	.75 (-.06, 1.59)
Now 35: at 80: 1 yr sick; recov	0.48 (0.27, 0.7)	0.38 (0.07, 0.69)	0.39 (0.08, 0.7)	0.33 (0.05, 0.63)	.46 (.13, .82)
Now 35: now: 5 yrs sick; recov	3.62 (2.17, 5.11)	2.62 (0.55, 4.7)	2.75 (0.65, 4.91)	5.24 (3.3, 7.3)	3.96 (1.64, 6.25)
Now 35: at 40: 5 yrs sick; recov	3.26 (1.94, 4.61)	2.34 (0.48, 4.22)	2.46 (0.58, 4.41)	4.67 (2.93, 6.54)	3.54 (1.44, 5.61)
Now 35: at 50: 5 yrs sick; recov	2.54 (1.52, 3.6)	1.84 (0.39, 3.3)	1.93 (0.45, 3.45)	3.56 (2.19, 5.02)	2.75 (1.13, 4.39)
Now 35: at 60: 5 yrs sick; recov	1.86 (1.12, 2.62)	1.41 (0.38, 2.45)	1.47 (0.41, 2.55)	2.47 (1.49, 3.5)	2.04 (.88, 3.22)
Now 35: at 70: 5 yrs sick; recov	1.19 (0.77, 1.63)	1.07 (0.46, 1.68)	1.12 (0.49, 1.74)	1.4 (0.82, 2)	1.43 (.74, 2.11)
Now 35: at 80: 5 yrs sick; recov	0.52 (0.4, 0.66)	0.82 (0.62, 1.03)	0.86 (0.66, 1.08)	0.31 (0.12, 0.5)	.9 (.66, 1.17)
Now 35: now: 1 yr sick; then die	5.01 (4.25, 5.82)	5.33 (3.42, 7.5)	5.39 (3.28, 7.66)	8.84 (6.62, 11.61)	6.4 (4.31, 8.69)
Now 35: at 40: 1 yr sick; then die	4.29 (3.59, 5.03)	5.18 (3.68, 6.95)	5.25 (3.57, 7.06)	6.1 (4.44, 8.05)	5.85 (4.18, 7.63)
Now 35: at 50: 1 yr sick; then die	3.12 (2.54, 3.71)	4.78 (3.74, 5.93)	4.85 (3.76, 6.13)	1.86 (0.96, 2.81)	4.78 (3.69, 6.07)
Now 35: at 60: 1 yr sick; then die	2.18 (1.74, 2.64)	4.14 (3.26, 5.16)	4.21 (3.29, 5.26)	-0.85 (-1.79, 0.05)	3.7 (2.77, 4.72)
Now 35: at 70: 1 yr sick; then die	1.37 (1.08, 1.68)	3.16 (2.35, 4.04)	3.22 (2.4, 4.13)	-2.05 (-3.13, -1.09)	2.55 (1.72, 3.44)
Now 35: at 80: 1 yr sick; then die	0.61 (0.48, 0.75)	1.66 (1.17, 2.2)	1.7 (1.2, 2.23)	-1.58 (-2.32, -0.96)	1.24 (.73, 1.77)
Now 35: now: 5 yrs sick; then die	5.39 (4.62, 6.18)	5.42 (3.14, 7.97)	5.6 (3.25, 8.08)	10.56 (7.89, 13.94)	5.87 (3.7, 8.43)
Now 35: at 40: 5 yrs sick; then die	4.67 (3.98, 5.39)	5.35 (3.58, 7.33)	5.53 (3.68, 7.48)	7.61 (5.7, 9.98)	5.52 (3.82, 7.48)
Now 35: at 50: 5 yrs sick; then die	3.43 (2.89, 4.01)	5.01 (3.89, 6.18)	5.16 (4.1, 6.43)	2.97 (2.01, 3.97)	4.72 (3.65, 5.96)
Now 35: at 60: 5 yrs sick; then die	2.38 (1.99, 2.81)	4.31 (3.47, 5.28)	4.43 (3.57, 5.44)	n/a	3.77 (2.93, 4.72)
Now 35: at 70: 5 yrs sick; then die	1.45 (1.2, 1.71)	3.13 (2.44, 3.91)	3.22 (2.52, 4.02)	-1.23 (-2.09, -0.47)	2.6 (1.89, 3.37)

Now 35: at 80: 5 yrs sick; then die	0.57 (0.46, 0.68)	1.28 (0.99, 1.59)	1.32 (1.04, 1.64)	-0.4 (-0.75, -0.1)	1.14 (.83, 1.48)
Now 65: Sudden death now	3.97 (3.22, 4.77)	3.55 (1.57, 5.53)	3.56 (1.63, 5.55)	5.28 (3.15, 7.63)	2.77 (.92, 4.64)
Now 65: Sudden death at 70	3.26 (2.6, 3.98)	3.09 (1.74, 4.51)	3.09 (1.72, 4.47)	0.91 (-0.7, 2.4)	2.35 (.95, 3.75)
Now 65: Sudden death at 80	1.98 (1.52, 2.47)	2.17 (1.25, 3.07)	2.17 (1.24, 3.07)	-5.5 (-7.73, -3.75)	1.59 (.55, 2.63)
Now 65: Sudden death at 90	0.46 (0.33, 0.59)	0.64 (0.09, 1.16)	0.64 (0.11, 1.17)	-4.07 (-5.53, -2.97)	.46 (-.14, 1.03)
Now 65: now: 1 yr sick; recov	2.85 (1.5, 4.24)	4.21 (2.33, 6.09)	4.37 (2.6, 6.27)	7.28 (5.11, 9.89)	7.46 (4.49, 10.79)
Now 65: at 70: 1 yr sick; recov	2.5 (1.32, 3.72)	3.7 (2.06, 5.32)	3.84 (2.27, 5.49)	6.02 (4.12, 8.16)	6.51 (3.9, 9.42)
Now 65: at 80: 1 yr sick; recov	1.72 (0.96, 2.51)	2.54 (1.49, 3.59)	2.63 (1.58, 3.71)	3.07 (1.87, 4.39)	4.33 (2.59, 6.24)
Now 65: now: 5 yrs sick; recov	3.54 (2.27, 4.86)	5.04 (3.34, 6.8)	5.28 (3.63, 7.01)	7.83 (5.85, 10.32)	8.39 (5.53, 11.52)
Now 65: at 70: 5 yrs sick; recov	3.09 (2.01, 4.22)	4.4 (2.98, 5.9)	4.6 (3.21, 6.06)	6.37 (4.68, 8.42)	7.25 (4.85, 9.86)
Now 65: at 80: 5 yrs sick; recov	2.04 (1.45, 2.66)	2.87 (2.05, 3.65)	3 (2.25, 3.84)	2.68 (1.83, 3.62)	4.45 (3.13, 5.91)
Now 65: now: 1 yr sick; then die	4.41 (3.72, 5.15)	1.52 (-0.37, 3.29)	1.54 (-0.34, 3.43)	7.46 (5.31, 9.98)	1.71 (-.03, 3.43)
Now 65: at 70: 1 yr sick; then die	3.62 (3.01, 4.28)	1.69 (0.5, 2.82)	1.71 (0.51, 2.91)	2.49 (1.17, 3.8)	1.69 (.58, 2.86)
Now 65: at 80: 1 yr sick; then die	2.18 (1.77, 2.62)	1.67 (0.9, 2.44)	1.68 (0.88, 2.43)	-4.66 (-6.57, -3.15)	1.48 (.6, 2.37)
Now 65: at 90: 1 yr sick; then die	0.45 (0.37, 0.54)	0.63 (0.32, 0.93)	0.65 (0.36, 0.95)	-2.09 (-2.89, -1.48)	.63 (.29, .98)
Now 65: now: 5 yrs sick; then die	4.73 (4.03, 5.46)	-1.04 (-3.13, 0.82)	-1 (-3.12, 1.13)	8.09 (5.86, 10.81)	.46 (-1.6, 2.6)
Now 65: at 70: 5 yrs sick; then die	3.92 (3.31, 4.55)	-0.14 (-1.47, 1.02)	-0.09 (-1.45, 1.28)	2.98 (1.82, 4.34)	.95 (-.47, 2.37)
Now 65: at 80: 5 yrs sick; then die	2.3 (1.91, 2.72)	1.16 (0.42, 1.84)	1.22 (0.51, 1.9)	-3.6 (-5.22, -2.34)	1.62 (.71, 2.51)
Age profiles					
Now 25: at 25: sudden death	4.63 (3.81, 5.51)	4.21 (-0.4, 9.04)	4.23 (-0.78, 9.1)	4.37 (-0.04, 8.82)	6.55 (1.83, 11.47)
Now 30: at 30: sudden death	4.59 (3.78, 5.47)	4.88 (1.73, 8.18)	4.88 (1.44, 8.31)	5.73 (2.57, 8.84)	6.73 (3.49, 10.24)
Now 35: at 35: sudden death	4.55 (3.74, 5.42)	5.3 (3.06, 7.79)	5.29 (2.84, 7.83)	6.75 (4.32, 9.34)	6.69 (4.28, 9.27)
Now 40: at 40: sudden death	4.49 (3.69, 5.36)	5.49 (3.68, 7.46)	5.47 (3.44, 7.66)	7.36 (5.16, 9.76)	6.43 (4.51, 8.53)
Now 45: at 45: sudden death	4.41 (3.61, 5.27)	5.44 (3.67, 7.33)	5.42 (3.49, 7.42)	7.56 (5.41, 10.02)	5.96 (4.04, 8.11)
Now 50: at 50: sudden	4.33	5.19	5.17	7.46	5.33

death	(3.54, 5.18)	(3.39, 7.12)	(3.27, 7.27)	(5.3, 9.97)	(3.39, 7.46)
Now 55: at 55: sudden death	4.23 (3.46, 5.07)	4.76 (2.95, 6.68)	4.75 (2.9, 6.78)	7.03 (4.87, 9.51)	4.56 (2.64, 6.56)
Now 60: at 60: sudden death	4.11 (3.35, 4.94)	4.2 (2.39, 6)	4.19 (2.4, 6.09)	6.29 (4.16, 8.66)	3.68 (1.93, 5.55)
Now 65: at 65: sudden death	3.97 (3.22, 4.77)	3.55 (1.57, 5.53)	3.56 (1.63, 5.55)	5.28 (3.15, 7.63)	2.77 (.92, 4.64)
Now 70: at 70: sudden death	3.78 (3.05, 4.57)	2.93 (0.5, 5.27)	2.95 (0.57, 5.34)	4.11 (1.8, 6.52)	1.92 (-.34, 4.17)
Now 75: at 75: sudden death	3.61 (2.9, 4.37)	2.43 (-0.48, 5.43)	2.47 (-0.48, 5.53)	3.22 (0.33, 5.98)	1.25 (-1.78, 4.45)
Now 80: at 80: sudden death	3.48 (2.79, 4.22)	2.09 (-1.85, 5.94)	2.14 (-1.77, 6.14)	2.8 (-0.78, 6.21)	.76 (-3.14, 4.8)
Now 25: at 35: 5 yrs sick then die	4.13 (3.5, 4.8)	5.2 (2.47, 8.23)	5.42 (2.69, 8.43)	5.3 (2.73, 7.92)	5.29 (2.61, 8.2)
Now 30: at 40: 5 yrs sick then die	4.07 (3.45, 4.73)	5.35 (3.48, 7.44)	5.54 (3.66, 7.67)	5.3 (3.48, 7.31)	5.32 (3.49, 7.35)
Now 35: at 45: 5 yrs sick then die	4.02 (3.41, 4.67)	5.23 (3.87, 6.71)	5.39 (4.02, 6.89)	5.09 (3.75, 6.65)	5.14 (3.83, 6.61)
Now 40: at 50: 5 yrs sick then die	3.94 (3.33, 4.57)	4.82 (3.72, 6.16)	4.96 (3.87, 6.21)	4.55 (3.42, 5.89)	4.73 (3.61, 5.93)
Now 45: at 55: 5 yrs sick then die	3.82 (3.23, 4.44)	4.17 (3.2, 5.31)	4.28 (3.27, 5.47)	3.7 (2.72, 4.82)	4.14 (3.05, 5.32)
Now 50: at 60: 5 yrs sick then die	3.71 (3.13, 4.31)	3.34 (2.45, 4.38)	3.43 (2.45, 4.53)	2.71 (1.82, 3.74)	3.43 (2.42, 4.59)
Now 55: at 65: 5 yrs sick then die	3.56 (3, 4.14)	2.38 (1.55, 3.31)	2.46 (1.54, 3.44)	1.52 (0.66, 2.43)	2.65 (1.68, 3.69)
Now 60: at 70: 5 yrs sick then die	3.37 (2.83, 3.93)	1.41 (0.63, 2.15)	1.48 (0.62, 2.32)	0.2 (-0.73, 1.05)	1.9 (.93, 2.92)
Now 65: at 75: 5 yrs sick then die	3.12 (2.62, 3.64)	0.61 (-0.27, 1.38)	0.66 (-0.2, 1.5)	-1.12 (-2.22, -0.17)	1.34 (.29, 2.4)
Now 70: at 80: 5 yrs sick then die	2.78 (2.31, 3.26)	0.25 (-0.73, 1.19)	0.3 (-0.66, 1.2)	-2.17 (-3.58, -1.05)	1.21 (.05, 2.36)
Now 75: at 85: 5 yrs sick then die	2.4 (1.98, 2.84)	0.45 (-0.71, 1.55)	0.5 (-0.66, 1.54)	-2.26 (-3.79, -0.97)	1.58 (.12, 2.96)
Now 80: at 90: 5 yrs sick then die	2.05 (1.67, 2.45)	1.04 (-0.39, 2.38)	1.1 (-0.29, 2.45)	-1.35 (-2.96, -0.02)	2.25 (.61, 3.8)
Now 25: ill 6 mo die 6 mo early	0.07 (0.06, 0.09)	0.36 (0.23, 0.51)	0.37 (0.23, 0.53)	0.15 (0.03, 0.27)	.3 (.16, .43)
Now 30: ill 6 mo die 6 mo early	0.09 (0.08, 0.11)	0.35 (0.24, 0.47)	0.36 (0.25, 0.49)	-0.05 (-0.16, 0.06)	.28 (.16, .39)
Now 35: ill 6 mo die 6 mo early	0.11 (0.09, 0.13)	0.32 (0.22, 0.42)	0.33 (0.23, 0.43)	-0.3 (-0.44, -0.18)	.24 (.14, .34)
Now 40: ill 6 mo die 6 mo early	0.14 (0.12, 0.17)	0.3 (0.19, 0.4)	0.31 (0.2, 0.42)	-0.61 (-0.83, -0.44)	.21 (.11, .33)
Now 45: ill 6 mo die 6 mo early	0.18 (0.15, 0.21)	0.28 (0.15, 0.42)	0.29 (0.15, 0.43)	-0.99 (-1.32, -0.73)	.2 (.05, .34)
Now 50: ill 6 mo die 6 mo early	0.22 (0.18, 0.26)	0.27 (0.09, 0.45)	0.28 (0.1, 0.46)	-1.37 (-1.83, -1.01)	.19 (-.01, .37)
Now 55: ill 6 mo die 6 mo early	0.26 (0.21, 0.31)	0.29 (0.07, 0.49)	0.3 (0.09, 0.52)	-1.73 (-2.33, -1.28)	.21 (-.02, .46)
Now 60: ill 6 mo die 6 mo early	0.31 (0.25, 0.37)	0.37 (0.1, 0.62)	0.38 (0.12, 0.65)	-2.06 (-2.78, -1.51)	.3 (.01, .58)

Now 65: ill 6 mo die 6 mo early	0.37 (0.3, 0.44)	0.52 (0.2, 0.83)	0.53 (0.22, 0.85)	-2.29 (-3.13, -1.65)	.47 (.12, .82)
Now 70: ill 6 mo die 6 mo early	0.44 (0.36, 0.53)	0.78 (0.34, 1.22)	0.79 (0.36, 1.22)	-2.37 (-3.31, -1.64)	.75 (.27, 1.2)
Now 75: ill 6 mo die 6 mo early	0.5 (0.41, 0.6)	1.14 (0.5, 1.77)	1.14 (0.5, 1.81)	-2.15 (-3.2, -1.26)	1.14 (.46, 1.77)
Now 80: ill 6 mo die 6 mo early	0.55 (0.45, 0.65)	1.59 (0.65, 2.52)	1.59 (0.67, 2.55)	-1.59 (-2.78, -0.6)	1.62 (.66, 2.49)
Income effects					
Now 45: Sudden death at 45 for income $10K	1.65 (1.29, 2.03)	2.11 (1.31, 2.96)	2.1 (1.23, 3)	3.06 (2.1, 4.16)	2.35 (1.48, 3.31)
Now 45: Sudden death at 45 for income $20K	2.62 (2.1, 3.19)	3.3 (2.13, 4.55)	3.29 (2.01, 4.62)	4.7 (3.29, 6.33)	3.65 (2.38, 5.07)
Now 45: Sudden death at 45 for income $30K	3.47 (2.81, 4.19)	4.32 (2.86, 5.89)	4.31 (2.71, 5.97)	6.08 (4.31, 8.12)	4.76 (3.17, 6.54)
Now 45: Sudden death at 45 for income $40K	4.26 (3.48, 5.09)	5.26 (3.54, 7.1)	5.24 (3.37, 7.19)	7.32 (5.23, 9.72)	5.77 (3.9, 7.86)
Now 45: Sudden death at 45 for income $50K	5 (4.13, 5.95)	6.13 (4.18, 8.22)	6.11 (3.99, 8.32)	8.47 (6.11, 11.18)	6.71 (4.6, 9.07)
Now 45: Sudden death at 45 for income $60K	5.71 (4.74, 6.76)	6.96 (4.81, 9.27)	6.94 (4.59, 9.39)	9.55 (6.93, 12.56)	7.6 (5.26, 10.22)
Now 45: Sudden death at 45 for income $70K	6.4 (5.34, 7.54)	7.76 (5.41, 10.28)	7.73 (5.18, 10.4)	10.58 (7.73, 13.86)	8.46 (5.91, 11.31)
Now 45: Sudden death at 45 for income $80K	7.07 (5.93, 8.3)	8.53 (6.01, 11.24)	8.51 (5.76, 11.38)	11.57 (8.5, 15.1)	9.29 (6.54, 12.35)
Now 45: Sudden death at 45 for income $90K	7.72 (6.51, 9.03)	9.29 (6.59, 12.18)	9.26 (6.32, 12.32)	12.53 (9.25, 16.3)	10.09 (7.16, 13.36)
Now 45: Sudden death at 45 for income $100K	8.36 (7.08, 9.75)	10.02 (7.16, 13.08)	9.99 (6.88, 13.24)	13.46 (9.98, 17.45)	10.87 (7.77, 14.34)
Now 45: Sudden death at 45 for income $110K	8.99 (7.64, 10.45)	10.74 (7.73, 13.97)	10.7 (7.43, 14.13)	14.36 (10.7, 18.58)	11.63 (8.36, 15.29)
Now 45: Sudden death at 45 for income $120K	9.61 (8.19, 11.15)	11.44 (8.28, 14.83)	11.41 (7.97, 15)	15.25 (11.4, 19.67)	12.38 (8.95, 16.22)
Now 45: Sudden death at 45 for income $130K	10.22 (8.74, 11.83)	12.14 (8.83, 15.68)	12.1 (8.51, 15.86)	16.11 (12.09, 20.74)	13.12 (9.53, 17.13)
Now 45: Sudden death at 45 for income $140K	10.82 (9.28, 12.5)	12.82 (9.38, 16.51)	12.78 (9.04, 16.7)	16.96 (12.78, 21.78)	13.84 (10.11, 18.02)
Now 45: Sudden death at 45 for income $150K	11.42 (9.81, 13.16)	13.49 (9.92, 17.33)	13.45 (9.56, 17.52)	17.8 (13.45, 22.8)	14.56 (10.68, 18.9)
Now 45: Sudden death at 45 for income $160K	12.01 (10.34, 13.81)	14.16 (10.45, 18.14)	14.12 (10.09, 18.34)	18.62 (14.11, 23.81)	15.26 (11.24, 19.76)

[a] These estimates are based on an arbitrarily specified initial risk of 0.004 and a risk reduction to 0.001. The risk reduction is thus of magnitude 0.003 and this average WTP for a microrisk reduction is calculated across these 3000 microrisks. The slight curvature of the utility function means that the average WTP amounts will differ somewhat with the size of the risk reduction over which they are calculated.

Table 5-7 Controlling for status quo effects: estimated indirect utility parameters
(See Section 5.6.5. 1,801 individuals, 7,520 completed choice sets, 22,560 alternatives; three types of exclusion criteria, limited set of scenario adjustment/exclusion controls.)

		Basic Model	With status quo effect
β_0	$\left(Y_i - c_i^j\right)^{(0.45)} cterm_i^j - \left(Y_i\right)^{(0.45)} yterm_i^j$	0.01015 (9.05)***	.01141 (9.86)***
α_{10}	$\Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$	-74.86 (7.05)***	-58.35 (5.28)***
α_{20}	$\Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	62.03 (1.85)*	67.25 (2.00)**
α_{21}	$\ldots\ age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	-2.119 (2.98)***	-2.159 (3.02)***
α_{30}	$\Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-608 (3.33)***	-549.3 (2.99)***
α_{31}	$\ldots\ age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	21.09 (2.85)***	20.38 (2.74)***
α_{32}	$\ldots\ age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-0.1975 (2.79)***	-.1917 (2.70)***
α_{40}	$\Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	196.2 (2.31)**	176.1 (2.07)**
α_{41}	$\ldots\ age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	-7.895 (2.29)**	-7.496 (2.17)**
α_{42}	$\ldots\ age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	0.07865 (2.35)**	.07512 (2.24)**
α_{50}	$\Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	127 (1.69)*	113.3 (1.51)
α_{51}	$\ldots\ age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	-4.42 -1.51	-4.316 (1.47)
α_{52}	$\ldots\ age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	0.05009 (1.81)*	.04901 (1.77)*
	Status quo effect:		
δ	1(neither program) [a]	-	-.234 (5.52)***
	Systematic selection correction term:		
α_{13}	$\ldots \left[P(sel_i) - \overline{P}\right] \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right]$	3.287 (2.25)**	3.293 (2.22)**
	Scenario adjustment variables [b]		
	$age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right) \times benefit\_never_i^j$	21.59 (5.35)***	22.36 (5.46)***
	$age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right) \times benefit\_never_i^j$	-0.2606	-.2702

	(4.45)***	(4.56)***
$age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right) \times benefit\_never_i^j$	0.129	.1338
	(3.94)***	(4.03)***
$\Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2 \times benefit\_never_i^j$	682.7	707.1
	(6.36)***	(6.46)***
$age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2 \times benefit\_never_i^j$	-12.74	-13.21
	(5.62)***	(5.73)***
$age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right]$	-25.49	-26.54
$\times \left[\log\left(pdvl_i^j + 1\right)\right] \times benefit\_never_i^j$	(5.13)***	(5.27)***
$age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right]$	0.3515	.3662
$\times \left[\log\left(pdvl_i^j + 1\right)\right] \times benefit\_never_i^j$	(4.56)***	(4.69)***
$age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right) \times \log\left(\left\|LEdiff < 0\right\| + 1\right)$	0.00308	.00301
	(2.27)**	(2.20)**
$age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right) \times \log\left(\left\|LEdiff < 0\right\| + 1\right)$	0.3589	.3672
	(2.19)**	(2.23)**
Max LogL	-11486.61	-11471.368
Alternatives	22560	22560

Absolute value of z statistics in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%
[a] We use a status quo indicator, rather than an "any program" indicator because we wish to simulate the indirect utility difference function in the case where the status quo effect is zero..

Table 5-8 Influence of status quo effects: average WTP for microrisk reductions
(See Section 5.6.5. Based on the parameter estimates in Table 5-7.)

	Basic Model	Net of status quo effects
Income = $42,000		
Now 45: Sudden death at 45	$ 7.59 (5.26, 9.99)	$ 5.95 (3.98, 7.94)
Now 45: at 45: 1 yr sick; recov	5.24 (2.86, 7.72)	3.90 (1.77, 5.98)
Now 45: at 45: 5 yrs sick; recov	7.44 (5.14, 9.87)	5.42 (3.44, 7.47)
Now 45: at 45: 1 yr sick; then die	7.11 (5.03, 9.51)	5.84 (4.03, 7.83)
Now 45: at 45: 5 yrs sick; then die	6.32 (4.15, 8.92)	5.40 (3.47, 7.43)
Income= $25,000		
Now 45: Sudden death at 45	5.44 (3.69, 7.24)	4.21 (2.74, 5.7)
Now 45: at 45: 1 yr sick; recov	3.93 (2.15, 5.79)	2.93 (1.33, 4.49)
Now 45: at 45: 5 yrs sick; recov	5.58 (3.86, 7.39)	4.06 (2.58, 5.6)
Now 45: at 45: 1 yr sick; then die	5.1 (3.54, 6.89)	4.14 (2.79, 5.63)
Now 45: at 45: 5 yrs sick; then die	4.55 (2.93, 6.5)	3.86 (2.42, 5.38)
Income= $67,500		
Now 45: Sudden death at 45	10.37 (7.33, 13.5)	8.23 (5.68, 10.82)
Now 45: at 45: 1 yr sick; recov	6.81 (3.71, 10.04)	5.06 (2.30, 7.77)
Now 45: at 45: 5 yrs sick; recov	9.67 (6.68, 12.84)	7.04 (4.47, 9.71)
Now 45: at 45: 1 yr sick; then die	9.72 (7.01, 12.84)	8.06 (5.71, 10.65)
Now 45: at 45: 5 yrs sick; then die	8.59 (5.77, 11.97)	7.39 (4.88, 10.03)
Latency (income=$42K)		
Now 35: Sudden death now	8.3 (5.52, 11.59)	6.70 (4.19, 9.09)
Now 35: Sudden death at 40	7.74 (5.4, 10.39)	6.01 (3.97, 8)
Now 35: Sudden death at 50	6.64 (5.04, 8.56)	4.80 (3.42, 6.24)

Now 35: Sudden death at 60	5.44 (4.22, 6.82)	3.67 (2.6, 4.85)
Now 35: Sudden death at 70	3.98 (2.98, 5.09)	2.53 (1.61, 3.49)
Now 35: Sudden death at 80	2.08 (1.45, 2.77)	1.25 (0.67, 1.85)
Now 35: now: 1 yr sick; recov	3.23 (0.64, 5.92)	2.10 (-0.23, 4.46)
Now 35: at 40: 1 yr sick; recov	2.84 (0.48, 5.31)	1.84 (-0.31, 4.00)
Now 35: at 50: 1 yr sick; recov	2.16 (0.27, 4.14)	1.39 (-0.33, 3.11)
Now 35: at 60: 1 yr sick; recov	1.61 (0.18, 3.1)	1.02 (-0.27, 2.31)
Now 35: at 70: 1 yr sick; recov	1.14 (0.22, 2.1)	0.71 (-0.11, 1.57)
Now 35: at 80: 1 yr sick; recov	0.72 (0.35, 1.12)	0.45 (0.12, 0.80)
Now 35: now: 5 yrs sick; recov	5.64 (3.21, 8.22)	3.78 (1.64, 6.02)
Now 35: at 40: 5 yrs sick; recov	5.05 (2.85, 7.4)	3.38 (1.46, 5.41)
Now 35: at 50: 5 yrs sick; recov	3.93 (2.21, 5.79)	2.63 (1.13, 4.2)
Now 35: at 60: 5 yrs sick; recov	2.93 (1.7, 4.27)	1.95 (0.87, 3.08)
Now 35: at 70: 5 yrs sick; recov	2.07 (1.34, 2.87)	1.37 (0.74, 2.04)
Now 35: at 80: 5 yrs sick; recov	1.32 (1.04, 1.62)	0.87 (0.62, 1.13)
Now 35: now: 1 yr sick; then die	7.57 (5.24, 10.36)	6.37 (4.27, 8.58)
Now 35: at 40: 1 yr sick; then die	7.29 (5.37, 9.59)	5.83 (4.12, 7.59)
Now 35: at 50: 1 yr sick; then die	6.52 (5.15, 8.18)	4.78 (3.61, 6.02)
Now 35: at 60: 1 yr sick; then die	5.45 (4.37, 6.74)	3.72 (2.73, 4.76)
Now 35: at 70: 1 yr sick; then die	4.01 (3.05, 5.07)	2.57 (1.7, 3.48)
Now 35: at 80: 1 yr sick; then die	2.05 (1.48, 2.69)	1.25 (0.72, 1.78)
Now 35: now: 5 yrs sick; then die	6.6 (3.94, 9.44)	5.79 (3.55, 8.29)
Now 35: at 40: 5 yrs sick; then die	6.63 (4.59, 8.9)	5.45 (3.73, 7.4)

Now 35: at 50: 5 yrs sick; then die	6.32 (4.96, 7.86)	4.68 (3.53, 5.87)
Now 35: at 60: 5 yrs sick; then die	5.45 (4.39, 6.66)	3.75 (2.82, 4.7)
Now 35: at 70: 5 yrs sick; then die	3.98 (3.14, 4.9)	2.59 (1.82, 3.34)
Now 35: at 80: 5 yrs sick; then die	1.75 (1.4, 2.14)	1.13 (0.8, 1.46)
Now 65: Sudden death now	4.07 (1.92, 6.34)	2.79 (0.79, 4.71)
Now 65: Sudden death at 70	3.77 (2.21, 5.47)	2.37 (0.91, 3.77)
Now 65: Sudden death at 80	3.02 (2.07, 3.96)	1.61 (0.65, 2.61)
Now 65: Sudden death at 90	1.04 (0.47, 1.6)	0.47 (-0.10, 1.03)
Now 65: now: 1 yr sick; recov	9.02 (5.68, 12.28)	7.27 (4.52, 10.24)
Now 65: at 70: 1 yr sick; recov	7.9 (4.97, 10.74)	6.35 (3.93, 8.95)
Now 65: at 80: 1 yr sick; recov	5.36 (3.48, 7.19)	4.23 (2.64, 5.90)
Now 65: now: 5 yrs sick; recov	10.42 (7.49, 13.5)	8.13 (5.56, 10.85)
Now 65: at 70: 5 yrs sick; recov	9.06 (6.55, 11.69)	7.03 (4.84, 9.35)
Now 65: at 80: 5 yrs sick; recov	5.77 (4.33, 7.22)	4.32 (3.10, 5.59)
Now 65: now: 1 yr sick; then die	2.46 (0.48, 4.52)	1.67 (-0.23, 3.43)
Now 65: at 70: 1 yr sick; then die	2.76 (1.46, 4.12)	1.66 (0.42, 2.88)
Now 65: at 80: 1 yr sick; then die	2.81 (2.04, 3.63)	1.47 (0.64, 2.34)
Now 65: at 90: 1 yr sick; then die	1.11 (0.79, 1.44)	0.62 (0.29, 0.97)
Now 65: now: 5 yrs sick; then die	0.69 (-1.47, 2.92)	0.32 (-1.76, 2.34)
Now 65: at 70: 5 yrs sick; then die	1.68 (0.18, 3.18)	0.83 (-0.65, 2.16)
Now 65: at 80: 5 yrs sick; then die	2.87 (2.02, 3.75)	1.55 (0.66, 2.40)

Age profiles

Now 25: at 25: sudden death	7.97 (2.58, 13.9)	6.62 (1.94, 11.41)
Now 30: at 30: sudden death	8.26 (4.34, 12.33)	6.76 (3.36, 10.10)
Now 35: at 35: sudden death	8.3 (5.52, 11.59)	6.7 (4.19, 9.09)
Now 40: at 40: sudden death	8.07 (5.73, 10.76)	6.42 (4.33, 8.46)
Now 45: at 45: sudden death	7.59 (5.26, 9.99)	5.95 (3.98, 7.94)
Now 50: at 50: sudden death	6.91 (4.56, 9.28)	5.32 (3.26, 7.24)
Now 55: at 55: sudden death	6.06 (3.92, 8.36)	4.55 (2.66, 6.45)
Now 60: at 60: sudden death	5.09 (3.05, 7.3)	3.68 (1.81, 5.52)
Now 65: at 65: sudden death	4.07 (1.92, 6.34)	2.79 (0.79, 4.71)
Now 70: at 70: sudden death	3.12 (0.57, 5.61)	1.97 (-0.42, 4.27)
Now 75: at 75: sudden death	2.37 (-0.99, 5.66)	1.32 (-1.82, 4.35)
Now 80: at 80: sudden death	1.79 (-2.54, 6.12)	0.85 (-3.24, 4.81)
Now 25: at 35: 5 yrs sick then die	6.64 (3.64, 9.81)	5.30 (2.74, 8.08)
Now 30: at 40: 5 yrs sick then die	6.72 (4.61, 9.12)	5.29 (3.49, 7.29)
Now 35: at 45: 5 yrs sick then die	6.54 (4.86, 8.34)	5.08 (3.71, 6.58)
Now 40: at 50: 5 yrs sick then die	6.11 (4.77, 7.7)	4.66 (3.55, 5.91)
Now 45: at 55: 5 yrs sick then die	5.47 (4.26, 6.94)	4.05 (3.02, 5.21)
Now 50: at 60: 5 yrs sick then die	4.69 (3.54, 6.05)	3.33 (2.35, 4.41)
Now 55: at 65: 5 yrs sick then die	3.84 (2.74, 5.04)	2.55 (1.60, 3.52)
Now 60: at 70: 5 yrs sick then die	3.02 (2, 4.09)	1.81 (0.86, 2.73)
Now 65: at 75: 5 yrs sick then die	2.43 (1.37, 3.48)	1.25 (0.18, 2.26)
Now 70: at 80: 5 yrs sick then die	2.34 (1.16, 3.51)	1.12 (-0.06, 2.29)
Now 75: at 85: 5 yrs sick then die	2.78 (1.35, 4.24)	1.49 (0.09, 2.85)
Now 80: at 90: 5 yrs sick then die	3.51 (1.82, 5.28)	2.15 (0.60, 3.68)

Now 25: ill 6 mo die 6 mo early	0.42 (0.27, 0.58)	0.30 (0.17, 0.44)
Now 30: ill 6 mo die 6 mo early	0.42 (0.29, 0.56)	0.28 (0.17, 0.40)
Now 35: ill 6 mo die 6 mo early	0.4 (0.29, 0.52)	0.24 (0.14, 0.34)
Now 40: ill 6 mo die 6 mo early	0.4 (0.29, 0.53)	0.22 (0.10, 0.32)
Now 45: ill 6 mo die 6 mo early	0.41 (0.27, 0.57)	0.2 (0.05, 0.34)
Now 50: ill 6 mo die 6 mo early	0.43 (0.25, 0.63)	0.19 (0.00, 0.37)
Now 55: ill 6 mo die 6 mo early	0.5 (0.27, 0.73)	0.21 (-0.01, 0.44)
Now 60: ill 6 mo die 6 mo early	0.64 (0.37, 0.91)	0.30 (0.03, 0.56)
Now 65: ill 6 mo die 6 mo early	0.89 (0.55, 1.22)	0.46 (0.13, 0.81)
Now 70: ill 6 mo die 6 mo early	1.26 (0.81, 1.73)	0.74 (0.30, 1.20)
Now 75: ill 6 mo die 6 mo early	1.75 (1.07, 2.42)	1.13 (0.50, 1.76)
Now 80: ill 6 mo die 6 mo early	2.33 (1.29, 3.38)	1.61 (0.73, 2.47)
Income effects		
Now 45: Sudden death at 45 for income $10K	3.07 (2.03, 4.14)	2.34 (1.46, 3.23)
Now 45: Sudden death at 45 for income $20K	4.72 (3.18, 6.31)	3.64 (2.34, 4.95)
Now 45: Sudden death at 45 for income $30K	6.11 (4.18, 8.1)	4.75 (3.12, 6.40)
Now 45: Sudden death at 45 for income $40K	7.35 (5.08, 9.69)	5.75 (3.84, 7.69)
Now 45: Sudden death at 45 for income $50K	8.5 (5.93, 11.15)	6.69 (4.53, 8.89)
Now 45: Sudden death at 45 for income $60K	9.59 (6.74, 12.52)	7.58 (5.19, 10.01)
Now 45: Sudden death at 45 for income $70K	10.62 (7.52, 13.82)	8.44 (5.83, 11.08)
Now 45: Sudden death at 45 for income $80K	11.61 (8.28, 15.06)	9.26 (6.46, 12.11)
Now 45: Sudden death at 45 for income $90K	12.57 (9.01, 16.25)	10.06 (7.07, 13.10)
Now 45: Sudden death at 45 for income $100K	13.51 (9.73, 17.4)	10.85 (7.67, 14.07)
Now 45: Sudden death at 45 for income $110K	14.41 (10.43, 18.52)	11.61 (8.27, 15.01)
Now 45: Sudden death at 45 for income $120K	15.3 (11.12, 19.61)	12.36 (8.85, 15.92)

Now 45: Sudden death at 45 for income $130K	16.17 (11.8, 20.68)	13.09 (9.43, 16.82)
Now 45: Sudden death at 45 for income $140K	17.02 (12.47, 21.72)	13.82 (10.00, 17.7)
Now 45: Sudden death at 45 for income $150K	17.86 (13.13, 22.74)	14.53 (10.56, 18.56)
Now 45: Sudden death at 45 for income $160K	18.68 (13.79, 23.74)	15.23 (11.12, 19.41)

Table 5-9 By discount rate assumption: estimated indirect utility parameters
(See Section 5.6.5; models include slightly different scenario rejection corrections; 2,407 individuals, 11,385 completed choice sets, 34,155 alternatives; one exclusion criterion, full set of scenario adjustment/exclusion controls from the main paper.)

		Common fixed discount rate assumptions		
		1 r =.03	**2** **r = .05**	3 r = .07
β_0	$\left[\frac{\left(Y_i - c_i^j\right)^{0.45} - 1}{0.45}\right] cterm_i^j - \left[\frac{\left(Y_i\right)^{0.45} - 1}{0.45}\right] yterm_i^j$	.01054 (8.48)***	.01344 (8.59)***	.01638 (8.66)***
α_{10}	$\Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$	-24.97 (2.25)**	-27.14 (2.22)**	-30.23 (2.24)**
α_{20}	$\Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	19.53 (0.77)	27.87 (0.87)	35.91 (0.89)
α_{21}	… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	-.9421 (1.96)*	-1.228 (2.07)**	-1.517 (2.07)**
α_{30}	$\Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-1150 (3.39)***	-1680 (4.03)***	-2093 (4.06)***
α_{31}	… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	46.68 (3.35)***	66.45 (3.91)***	81.47 (3.94)***
α_{32}	… $age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-.44 (3.20)***	-.6199 (3.72)***	-.756 (3.77)***
α_{40}	$\Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	443.2 (3.03)***	789.9 (3.69)***	1157 (3.73)***
α_{41}	… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	-18.41 (2.99)***	-31.6 (3.59)***	-45.38 (3.66)***
α_{42}	… $age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	.1725 (2.77)***	.2921 (3.35)***	.4166 (3.46)***
α_{50}	$\Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	-111.5 (2.90)***	-188.1 (3.18)***	-276.5 (3.12)***
α_{51}	… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	2.288 (3.02)***	3.657 (3.27)***	5.24 (3.23)***
	Status quo effect:			
	1(neither program)	-.05003 (0.84)	-.09183 (1.56)	-.1250 (2.17)**
	Scenario adjustment/rejection and systematic selection correction terms: [a]	Yes	**Yes**	Yes
	Total alternatives (= choices*3)	34155	34155	34155
	Max LogL	-14969.824	-14841.337	-14747.977

Absolute value of z statistics in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%
[a] For comparability, we constrain the set of base variables to be the same in each model as in the model for the 0.05 discount rate. The scenario adjustment/rejection and systematic selection correction terms are the same as in Cameron and DeShazo (2011)

Table 5-10 By discount rate assumption: average WTP for microrisk reductions
(See Section 5.6.5. Based on the parameter estimates in Table 5-9.)

	1 $r = 0.03$ [a]	**2** $r = 0.05$	3 $r = 0.07$
Income = $42,000			
Now 45: Sudden death at 45	$ 8.33 (4.45, 12.46)	$ 6.74 (3.12, 10.68)	$ 5.48 (1.44, 9.62)
Now 45: at 45: 1 yr sick; recov	2.58 (.4, 4.85)	2.42 (.51, 4.49)	2.25 (.54, 4.13)
Now 45: at 45: 5 yrs sick; recov	3.39 (1.13, 5.77)	3.05 (1.15, 5.07)	2.74 (.98, 4.55)
Now 45: at 45: 1 yr sick; then die	9.22 (5.58, 13.11)	8.09 (4.6, 11.82)	7.12 (3.26, 11.19)
Now 45: at 45: 5 yrs sick; then die	9.75 (5.84, 13.83)	9.09 (5.33, 13.44)	8.35 (4.28, 12.6)
Income= $25,000			
Now 45: Sudden death at 45	5.91 (3, 9.01)	4.81 (2.09, 7.76)	3.91 (.88, 7.02)
Now 45: at 45: 1 yr sick; recov	1.94 (.3, 3.65)	1.82 (.39, 3.37)	1.69 (.41, 3.1)
Now 45: at 45: 5 yrs sick; recov	2.55 (.85, 4.33)	2.29 (.86, 3.81)	2.06 (.74, 3.41)
Now 45: at 45: 1 yr sick; then die	6.59 (3.86, 9.51)	5.83 (3.21, 8.63)	5.16 (2.27, 8.2)
Now 45: at 45: 5 yrs sick; then die	7.04 (4.11, 10.09)	6.63 (3.81, 9.88)	6.13 (3.08, 9.3)
Income= $67,500			
Now 45: Sudden death at 45	11.5 (6.46, 16.88)	9.26 (4.56, 14.39)	7.51 (2.25, 12.9)
Now 45: at 45: 1 yr sick; recov	3.34 (.51, 6.3)	3.14 (.66, 5.83)	2.92 (.7, 5.37)
Now 45: at 45: 5 yrs sick; recov	4.4 (1.47, 7.49)	3.96 (1.48, 6.59)	3.56 (1.28, 5.91)
Now 45: at 45: 1 yr sick; then die	12.63 (7.9, 17.7)	10.99 (6.44, 15.85)	9.62 (4.59, 14.92)
Now 45: at 45: 5 yrs sick; then die	13.21 (8.14, 18.53)	12.19 (7.3, 17.86)	11.13 (5.83, 16.68)
Latency (income=$42K)			
Now 35: Sudden death now	4.74 (.18, 9.45)	.72 (-4.3, 5.92)	-1.77 (-7.72, 3.83)
Now 35: Sudden death at 40	4.39 (.35, 8.45)	1.17 (-2.79, 5.14)	-.43 (-4.73, 3.68)
Now 35: Sudden death at 50	3.78 (.73, 6.96)	1.91 (-.61, 4.44)	1.41 (-.72, 3.7)
Now 35: Sudden death at 60	3.23 (.88, 5.7)	2.3 (.32, 4.37)	2.09 (.5, 3.83)
Now 35: Sudden death at 70	2.63 (.43, 5.02)	2.19 (.43, 4)	1.79 (.55, 3.15)
Now 35: Sudden death at 80	1.73 (-.32, 3.77)	1.4 (.21, 2.63)	.87 (.25, 1.54)
Now 35: now: 1 yr sick; recov	2.49 (0, 5.14)	1.92 (-.29, 4.27)	1.57 (-.39, 3.66)

Now 35: at 40: 1 yr sick; recov	2.34 (-.01, 4.83)	1.74 (-.3, 3.87)	1.35 (-.37, 3.17)
Now 35: at 50: 1 yr sick; recov	2.04 (0, 4.17)	1.38 (-.26, 3.11)	.95 (-.28, 2.28)
Now 35: at 60: 1 yr sick; recov	1.73 (.05, 3.54)	1.06 (-.17, 2.35)	.63 (-.17, 1.5)
Now 35: at 70: 1 yr sick; recov	1.4 (.09, 2.77)	.77 (-.05, 1.62)	.38 (-.04, .85)
Now 35: at 80: 1 yr sick; recov	.97 (.17, 1.78)	.47 (.11, .86)	.2 (.05, .37)
Now 35: now: 5 yrs sick; recov	3.5 (1.03, 6.21)	2.68 (.5, 4.89)	2.16 (.23, 4.21)
Now 35: at 40: 5 yrs sick; recov	3.29 (1, 5.84)	2.42 (.44, 4.41)	1.86 (.21, 3.61)
Now 35: at 50: 5 yrs sick; recov	2.88 (.91, 5.04)	1.92 (.35, 3.49)	1.3 (.15, 2.5)
Now 35: at 60: 5 yrs sick; recov	2.44 (.83, 4.18)	1.45 (.34, 2.59)	.83 (.13, 1.57)
Now 35: at 70: 5 yrs sick; recov	1.96 (.76, 3.25)	1.01 (.33, 1.72)	.48 (.13, .84)
Now 35: at 80: 5 yrs sick; recov	1.32 (.65, 2.02)	.6 (.31, .9)	.25 (.13, .37)
Now 35: now: 1 yr sick; then die	7.64 (3.33, 12.29)	4.47 (-.13, 9.06)	2.61 (-2.62, 7.9)
Now 35: at 40: 1 yr sick; then die	6.86 (3.22, 10.79)	4.07 (.55, 7.65)	2.61 (-.96, 6.06)
Now 35: at 50: 1 yr sick; then die	5.54 (3.05, 8.22)	3.55 (1.52, 5.79)	2.73 (1.01, 4.65)
Now 35: at 60: 1 yr sick; then die	4.42 (2.46, 6.68)	3.13 (1.46, 5.01)	2.56 (1.12, 4.21)
Now 35: at 70: 1 yr sick; then die	3.36 (1.33, 5.56)	2.54 (.87, 4.28)	1.87 (.68, 3.18)
Now 35: at 80: 1 yr sick; then die	2.02 (.21, 3.87)	1.42 (.34, 2.53)	.81 (.26, 1.4)
Now 35: now: 5 yrs sick; then die	11.59 (6.57, 16.94)	9.66 (5.12, 14.86)	8.63 (3.45, 14.18)
Now 35: at 40: 5 yrs sick; then die	10.45 (6.25, 14.96)	8.49 (4.88, 12.61)	7.33 (3.94, 11.12)
Now 35: at 50: 5 yrs sick; then die	8.38 (5.57, 11.45)	6.48 (4.46, 8.88)	5.2 (3.48, 7.27)
Now 35: at 60: 5 yrs sick; then die	6.48 (4.53, 8.75)	4.77 (3.22, 6.6)	3.44 (2.03, 5.07)
Now 35: at 70: 5 yrs sick; then die	4.56 (2.75, 6.64)	3.1 (1.67, 4.67)	1.87 (.9, 3)
Now 35: at 80: 5 yrs sick; then die	1.98 (.82, 3.21)	1.04 (.47, 1.63)	.46 (.21, .74)
Now 65: Sudden death now	6.44 (1.66, 11.25)	5.91 (1.61, 10.24)	5.4 (1.33, 9.67)
Now 65: Sudden death at 70	5 (1.31, 8.73)	4.37 (1.4, 7.29)	3.81 (1.19, 6.58)
Now 65: Sudden death at 80	2.25 (-.07, 4.72)	1.73 (-.07, 3.49)	1.37 (-.22, 2.89)
Now 65: Sudden death at 90	-.13 (-1.79, 1.56)	-.05 (-1.16, 1.02)	.02 (-.77, .73)

Now 65: now: 1 yr sick; recov	4.06 (1.57, 6.7)	3.83 (1.7, 6.24)	3.63 (1.54, 5.83)
Now 65: at 70: 1 yr sick; recov	3.64 (1.38, 6.01)	3.33 (1.42, 5.44)	3.04 (1.29, 4.86)
Now 65: at 80: 1 yr sick; recov	2.55 (.8, 4.39)	2.14 (.86, 3.52)	1.75 (.73, 2.83)
Now 65: now: 5 yrs sick; recov	4.1 (1.74, 6.73)	3.85 (1.77, 6.07)	3.62 (1.72, 5.57)
Now 65: at 70: 5 yrs sick; recov	3.62 (1.52, 5.97)	3.28 (1.51, 5.16)	2.95 (1.41, 4.53)
Now 65: at 80: 5 yrs sick; recov	2.21 (.72, 3.84)	1.81 (.77, 2.9)	1.45 (.69, 2.28)
Now 65: now: 1 yr sick; then die	4.21 (-.52, 8.63)	3.67 (-.45, 8.06)	3.16 (-.92, 7.29)
Now 65: at 70: 1 yr sick; then die	3.19 (-.14, 6.5)	2.73 (-.04, 5.45)	2.33 (-.04, 4.74)
Now 65: at 80: 1 yr sick; then die	1.2 (-.73, 3.19)	.98 (-.51, 2.54)	.85 (-.52, 2.12)
Now 65: at 90: 1 yr sick; then die	-.05 (-1.09, 1.06)	.02 (-.64, .64)	.06 (-.39, .45)
Now 65: now: 5 yrs sick; then die	.55 (-4.47, 5.12)	.15 (-4.45, 4.52)	-.19 (-4.63, 4.34)
Now 65: at 70: 5 yrs sick; then die	.09 (-3.48, 3.34)	-.05 (-2.98, 2.9)	-.11 (-2.64, 2.47)
Now 65: at 80: 5 yrs sick; then die	-.55 (-2.42, 1.35)	-.28 (-1.78, 1.16)	-.02 (-1.24, 1.18)
Age profiles			
Now 25: at 25: sudden death	-3.28 (-13.56, 6.44)	-10.54 (-22.1, .32)	-14.57 (-28.12, -2.9)
Now 30: at 30: sudden death	1.31 (-5.34, 8.07)	-4.23 (-11.74, 3.3)	-7.44 (-16.61, .51)
Now 35: at 35: sudden death	4.74 (.18, 9.45)	.72 (-4.3, 5.92)	-1.77 (-7.72, 3.83)
Now 40: at 40: sudden death	7.06 (2.96, 11.26)	4.37 (.42, 8.38)	2.52 (-1.9, 6.96)
Now 45: at 45: sudden death	8.33 (4.45, 12.46)	6.74 (3.12, 10.68)	5.48 (1.44, 9.62)
Now 50: at 50: sudden death	8.78 (5, 12.88)	7.98 (4.37, 11.92)	7.16 (3.35, 11.21)
Now 55: at 55: sudden death	8.5 (4.53, 12.37)	8.15 (4.44, 12.2)	7.65 (3.73, 11.52)
Now 60: at 60: sudden death	7.64 (3.53, 11.56)	7.39 (3.59, 11.26)	7.01 (3.17, 10.71)
Now 65: at 65: sudden death	6.44 (1.66, 11.25)	5.91 (1.61, 10.24)	5.4 (1.33, 9.67)
Now 70: at 70: sudden death	5.2 (-.61, 11.21)	4 (-1.59, 9.24)	3.07 (-2.43, 8.32)
Now 75: at 75: sudden death	4.31 (-2.8, 11.97)	2.05 (-5.46, 8.98)	.36 (-7.09, 7.1)
Now 80: at 80: sudden death	3.83 (-5.33, 13.4)	.16 (-9.58, 9.16)	-2.63 (-12.75, 6.4)
Now 25: at 35: 5 yrs sick then die	10.27	8.31	7.93

	(4.05, 17.09)	(3.36, 13.66)	(3.92, 12.36)
Now 30: at 40: 5 yrs sick then die	10.12 (5.68, 14.99)	8.06 (4.54, 12.11)	7.13 (4.21, 10.52)
Now 35: at 45: 5 yrs sick then die	9.39 (5.89, 13.12)	7.43 (4.75, 10.48)	6.2 (3.91, 8.73)
Now 40: at 50: 5 yrs sick then die	8.09 (5.26, 11.15)	6.44 (4.25, 8.98)	5.16 (3.31, 7.26)
Now 45: at 55: 5 yrs sick then die	6.35 (3.96, 8.93)	5.12 (3.21, 7.34)	3.99 (2.36, 5.91)
Now 50: at 60: 5 yrs sick then die	4.45 (2.19, 6.71)	3.63 (1.84, 5.66)	2.78 (1.26, 4.47)
Now 55: at 65: 5 yrs sick then die	2.51 (.32, 4.75)	2.08 (.31, 3.9)	1.59 (.23, 3.14)
Now 60: at 70: 5 yrs sick then die	.8 (-1.37, 3.07)	.69 (-1.03, 2.44)	.56 (-.8, 2.04)
Now 65: at 75: 5 yrs sick then die	-.31 (-2.61, 1.98)	-.22 (-2.03, 1.58)	-.07 (-1.47, 1.42)
Now 70: at 80: 5 yrs sick then die	-.33 (-2.8, 2.19)	-.18 (-2.16, 1.69)	.06 (-1.57, 1.67)
Now 75: at 85: 5 yrs sick then die	.99 (-2.16, 4.21)	1.04 (-1.38, 3.41)	1.11 (-.96, 3.18)
Now 80: at 90: 5 yrs sick then die	3.07 (-.96, 7.52)	2.9 (-.23, 5.96)	2.6 (.06, 5.21)
Now 25: ill 6 mo die 6 mo early	1.37 (.6, 2.21)	.56 (.3, .85)	.19 (.1, .3)
Now 30: ill 6 mo die 6 mo early	.93 (.29, 1.58)	.44 (.2, .68)	.17 (.08, .27)
Now 35: ill 6 mo die 6 mo early	.47 (-.07, 1.01)	.26 (.05, .48)	.12 (.04, .21)
Now 40: ill 6 mo die 6 mo early	.05 (-.44, .57)	.08 (-.14, .31)	.06 (-.04, .15)
Now 45: ill 6 mo die 6 mo early	-.3 (-.87, .26)	-.11 (-.4, .16)	-.03 (-.16, .1)
Now 50: ill 6 mo die 6 mo early	-.55 (-1.24, .14)	-.27 (-.64, .06)	-.11 (-.3, .06)
Now 55: ill 6 mo die 6 mo early	-.63 (-1.39, .14)	-.35 (-.81, .05)	-.17 (-.42, .06)
Now 60: ill 6 mo die 6 mo early	-.49 (-1.39, .42)	-.29 (-.84, .2)	-.14 (-.46, .15)
Now 65: ill 6 mo die 6 mo early	-.09 (-1.17, 1)	-.02 (-.69, .63)	.03 (-.43, .43)
Now 70: ill 6 mo die 6 mo early	.64 (-.81, 2.2)	.55 (-.43, 1.51)	.45 (-.24, 1.09)
Now 75: ill 6 mo die 6 mo early	1.72 (-.32, 3.84)	1.48 (.07, 2.96)	1.17 (.13, 2.22)
Now 80: ill 6 mo die 6 mo early	3.15 (.33, 6.08)	2.74 (.68, 4.92)	2.19 (.62, 3.77)
Now 25: at 25: 1 year sick; recover	2.42 (-.5, 5.8)	1.28 (-1.45, 4.14)	0.79 (-1.63, 3.34)
Now 30: at 30: 1 year sick; recover	2.47 (-.17, 5.4)	1.63 (-.82, 4.2)	1.2 (-.92, 3.5)
Now 35: at 35: 1 year sick; recover	2.49 (0, 5.14)	1.92 (-.29, 4.27)	1.57 (-.39, 3.66)
Now 40: at 40: 1 year sick; recover	2.52 (.25, 4.91)	2.19 (.22, 4.37)	1.93 (.12, 3.93)
Now 45: at 45: 1 year sick; recover	2.58 (.4, 4.85)	2.42 (.51, 4.49)	2.25 (.54, 4.13)

Now 50: at 50: 1 year sick; recover	2.72 (.53, 5.01)	2.66 (.76, 4.67)	2.55 (.74, 4.42)
Now 55: at 55: 1 year sick; recover	2.99 (.72, 5.3)	2.94 (.96, 5.01)	2.86 (1.01, 4.81)
Now 60: at 60: 1 year sick; recover	3.42 (1.03, 5.86)	3.31 (1.27, 5.52)	3.2 (1.25, 5.22)
Now 65: at 65: 1 year sick; recover	4.06 (1.57, 6.7)	3.83 (1.7, 6.24)	3.63 (1.54, 5.83)
Now 70: at 70: 1 year sick; recover	4.94 (2.25, 7.82)	4.57 (2.22, 7.19)	4.21 (2.01, 6.6)
Now 75: at 75: 1 year sick; recover	6.15 (2.97, 9.51)	5.62 (3.05, 8.68)	5.06 (2.63, 7.7)
Now 80: at 80: 1 year sick; recover	7.71 (3.99, 11.72)	7 (3.99, 10.56)	6.2 (3.41, 9.29)
Income effects			
Now 45: Sudden death at 45 for income $10K	$ 3.29 (1.54, 5.15)	$ 2.70 (1.07, 4.46)	$ 2.20 (.38, 4.06)
Now 45: Sudden death at 45 for income $20K	5.11 (2.54, 7.85)	4.17 (1.77, 6.77)	3.4 (.72, 6.14)
Now 45: Sudden death at 45 for income $30K	6.66 (3.44, 10.09)	5.41 (2.41, 8.68)	4.4 (1.05, 7.84)
Now 45: Sudden death at 45 for income $40K	8.06 (4.28, 12.08)	6.53 (3.01, 10.36)	5.31 (1.37, 9.34)
Now 45: Sudden death at 45 for income $50K	9.37 (5.1, 13.92)	7.57 (3.58, 11.91)	6.14 (1.69, 10.71)
Now 45: Sudden death at 45 for income $60K	10.61 (5.88, 15.65)	8.55 (4.15, 13.36)	6.94 (2.01, 11.99)
Now 45: Sudden death at 45 for income $70K	11.8 (6.65, 17.28)	9.49 (4.69, 14.73)	7.69 (2.33, 13.2)
Now 45: Sudden death at 45 for income $80K	12.94 (7.4, 18.85)	10.4 (5.23, 16.04)	8.42 (2.65, 14.35)
Now 45: Sudden death at 45 for income $90K	14.05 (8.14, 20.36)	11.28 (5.76, 17.3)	9.12 (2.96, 15.46)
Now 45: Sudden death at 45 for income $100K	15.14 (8.87, 21.83)	12.13 (6.28, 18.51)	9.81 (3.28, 16.52)
Now 45: Sudden death at 45 for income $110K	16.2 (9.59, 23.25)	12.96 (6.8, 19.69)	10.48 (3.59, 17.55)
Now 45: Sudden death at 45 for income $120K	17.24 (10.3, 24.63)	13.78 (7.31, 20.84)	11.13 (3.91, 18.56)
Now 45: Sudden death at 45 for income $130K	18.25 (11.01, 25.99)	14.58 (7.82, 21.96)	11.77 (4.22, 19.53)
Now 45: Sudden death at 45 for income $140K	19.26 (11.71, 27.31)	15.36 (8.32, 23.05)	12.4 (4.53, 20.49)
Now 45: Sudden death at 45 for income $150K	20.24 (12.4, 28.61)	16.14 (8.82, 24.12)	13.01 (4.84, 21.42)
Now 45: Sudden death at 45 for income $160K	21.22 (13.09, 29.89)	16.9 (9.32, 25.18)	13.62 (5.16, 22.33)

[a] Different discounting assumptions mean that the present *discounted* future health states in the model must all be recalculated prior to the use of these variables to explain program choices. Different variables mean different estimates for the indirect utility parameters.

[b] These estimates are based on an arbitrarily specified initial risk of 0.004 and a risk reduction to 0.001. The risk reduction is thus of magnitude 0.003 and this average WTP for a microrisk reduction is calculated across these 3000 microrisks. The slight curvature of the utility function means that the average WTP amounts will differ somewhat with the size of the risk reduction over which they are calculated.

Table 5-11 Common 5% discount rate versus individual-specific discount rates

Column (1) reproduces the parameter estimates used in the main paper. Column (2) imposes calculated individual-specific financial discount rates based on the model estimated in Bosworth et al. (2011).

Parameter	Constructed Variable	(1) Estimated imposing r=0.05 for all	(2) Estimated using fitted individual r values[a]
β_0	$\left[\frac{\left(Y_i - c_i^j\right)^{0.45} - 1}{0.45}\right] cterm_i^j - \left[\frac{\left(Y_i\right)^{0.45} - 1}{0.45}\right] yterm_i^j$	.01344 (8.59)***	.01821 (8.29)***
α_{10}	$\Delta\pi_i^{jS} \log\left(pdvi_i^j + 1\right)$	-27.14 (2.22)**	-23 (1.63)
α_{20}	$\Delta\pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	27.87 (0.87)	51.49 (1.24)
α_{21}	$\ldots \; age_{i0} \times \Delta\pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	-1.228 (2.07)**	-1.846 (2.33)**
α_{30}	$\Delta\pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-1680 (4.03)***	-1799 (3.68)***
α_{31}	$\ldots \; age_{i0} \times \Delta\pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	66.45 (3.91)***	69.4 (3.46)***
α_{32}	$\ldots \; age_{i0}^2 \times \Delta\pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-.6199 (3.72)***	-.639 (3.23)***
α_{40}	$\Delta\pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	789.9 (3.69)***	970.9 (3.49)***
α_{41}	$\ldots \; age_{i0} \times \Delta\pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	-31.6 (3.59)***	-38.17 (3.27)***
α_{42}	$\ldots \; age_{i0}^2 \times \Delta\pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	.2921 (3.35)***	.3492 (2.98)***
α_{50}	$\Delta\pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	-188.1 (3.18)***	-232.3 (2.77)***
α_{51}	$\ldots$ $age_{i0} \times \Delta\pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	3.657 (3.27)***	4.137 (2.53)**
α_{51}	1(no program) = "status quo" indicator	-.09183 (1.56)	-.1193 (2.09)**
	Scenario adjustment/rejection controls (see main paper, Appendix Table A.1 for full variable names):[b]		
	Would never benefit?		
	b7term_bn	-.01002 (1.97)**	.01076 (1.42)
	dilog_bn	420.9 (8.18)***	432.6 (7.03)***
	dllog_agenow2_bn	.1406	.1636

	(4.01)***	(3.86)***
dllog2_bn	1200	1857
	(10.44)***	(10.83)***
dllog2_agenow_bn	-18.52	-28.27
	(7.19)***	(7.64)***
didllog_bn	-699.7	-829.9
	(4.45)***	(3.59)***
didllog_agenow_bn	6.885	6.082
	(2.57)**	(1.57)
Log(|pos. life expect. diff|+1):		
dllog_logldpos	672.9	662.6
	(2.62)***	(2.17)**
dllog_agenow_logldpos	-29.49	-27
	(2.67)***	(2.05)**
dllog_agenow2_logldpos	.2929	.2432
	(2.55)**	(1.77)*
dllog2_logldpos	-321.8	-359.7
	(2.41)**	(2.02)**
dllog2_agenow_logldpos	13.57	13.91
	(2.35)**	(1.78)*
dllog2_agenow2_logldpos	-.1313	-.1183
	(2.18)**	(1.44)
didllog_logldpos	104.7	122.9
	(2.75)***	(2.26)**
didllog_agenow_logldpos	-2.037	-2.258
	(2.63)***	(2.00)**
noprogram_logldpos	.09247	.07931
	(2.40)**	(2.23)**
Log(|neg. life expect. diff|+1):		
drlog_logldneg	12.5	16.21
	(1.65)*	(1.59)
dllog_logldneg	592.3	637.3
	(3.14)***	(2.89)***
dllog_agenow_logldneg	-26.03	-28.22
	(3.37)***	(3.08)***
dllog_agenow2_logldneg	.2591	.281
	(3.41)***	(3.09)***
dllog2_logldneg	-302.9	-388.1
	(3.14)***	(3.12)***
dllog2_agenow_logldneg	12.97	16.75
	(3.25)***	(3.17)***
dllog2_agenow2_logldneg	-.128	-.1658
	(3.23)***	(3.10)***
didllog_logldneg	56.41	60.73
	(2.13)**	(1.61)

didllog_agenow_logldneg	-.9817	-.9896
	(1.92)*	(1.32)
Shortens life most? Incorrect answer:		
dllog_shortwrong	247.6	237
	(2.91)***	(2.37)**
dllog_agenow_shortwrong	-5.352	-4.353
	(3.23)***	(2.21)**
dllog2_shortwrong	-131.2	-151.8
	(2.85)***	(2.57)**
dllog2_agenow_shortwrong	3.154	3.31
	(3.47)***	(2.74)***
noprogram_shortwrong	.099	.09696
	(2.36)**	(2.31)**
Failed risk comprehension test:		
dllog_nocomprisk	347.9	405.2
	(2.62)***	(2.29)**
dllog_agenow_nocomprisk	-11.98	-14.29
	(2.25)**	(2.03)**
dllog_agenow2_nocomprisk	.1365	.1644
	(2.74)***	(2.48)**
dllog2_agenow2_nocomprisk	-.01305	-.01665
	(2.12)**	(1.95)*
Status quo b/c reject scenario:		
b7term_reject	.6498	1.116
	(21.83)***	(21.08)***
Ingored affordability.:		
b7term_affordmiss	-.009186	-.01207
	(4.76)***	(4.41)***
dilog_affordmiss	-26.88	-45.07
	(1.75)*	(2.50)**
didllog_affordmiss	23.3	53.11
	(1.92)*	(2.96)***
noprogram_affordmiss	.1104	.1094
	(1.72)*	(1.72)*
Dev. from median select. prob:		
swrdilog	190	190
	(2.23)**	(1.80)*
Number of alternatives	34,155	34,155
Number of choices	11,385	11,385
Log L	-14841.337	-14773.197

[a] Fitted individual discount rates are computed as a function of observable individual characteristics used in Bosworth et al. (2011): discounti = exp(-0.0127*(subjective life expectancy) - 0.0381*(8.81) - 4.49*(age/100) + 0.432*(age/100)*(years of education) + 0.0157*female +0.522*female*nonwhite-0.185*(years of education) + 0.272 * (income<$27,500?) -0.396*(Hispanic)+ 0.675). The "subjective recovery likelihood variable, no collected in the survey used for this study, is set equal to the sample mean in the Bosworth, Cameron, and DeShazo study, since both samples are drawn from

the same population. Across the 2407 respondents to this survey, the mean calculated individual discount rate is 0.0839. The standard deviation is 0.0306. The minimum and maximum calculated individual discount rates are 0.0300 and 0.4817.

[b] For comparability, we assume the identical functional form for the scenario adjustment/rejection control variables in both specifications. The results are fairly robust.

Table 5-12 WTP based on individual discount rates vs. common 5% discount rate

(See Section 5.6.5. Based on the parameter estimates in Table 5-11. Column (1) imposes calculated individual-specific financial discount rates based on the model estimated in Bosworth et al. (2011). Column (2) uses the other fitted utility parameters from that model, but imposes a 3% discount rate for all respondents in simulating the WTP estimates; Column (3) and Column (4) impose 5% and 7% discount rates for all respondents; Column (5) contains the simulated WTP amounts from our main model in the paper which imposes a 5% discount rate for all respondents in the estimation phase, as well as the simulation of WTP. This column is provided for comparison with Column (3).

	(1) Estimated using fitted individual discount rates	(2) Simulated WTP based on r=0.03 for all	**(3) Simulated WTP based on r=0.05 for all**	(4) Simulated WTP based on r=0.07 for all	**(5) vs. Estimated imposing r=0.05 for all**
Income = $42,000					
	$ 3.31	$ 8.63	**$ 6.82**	$ 5.55	**$ 6.74**
Now 45: Sudden death at 45	(1, 5.61)	(2.64, 15.01)	**(2.2, 11.78)**	(1.81, 9.41)	**(3.12, 10.68)**
	1.51	2.27	**2.05**	1.88	**2.42**
Now 45: at 45: 1 yr sick; recov	(.18, 2.86)	(.25, 4.36)	**(.2, 3.92)**	(0.19, 3.59)	**(.51, 4.49)**
	1.63	2.87	**2.47**	2.2	**3.05**
Now 45: at 45: 5 yrs sick; recov	(.41, 2.82)	(.83, 4.94)	**(.67, 4.32)**	(0.56, 3.87)	**(1.15, 5.07)**
	4.41	10.59	**8.55**	7.09	**8.09**
Now 45: at 45: 1 yr sick; then die	(2.5, 6.65)	(4.78, 16.8)	**(4.12, 13.46)**	(3.58, 10.88)	**(4.6, 11.82)**
	4.54	12.68	**10.11**	8.19	**9.09**
Now 45: at 45: 5 yrs sick; then die	(2.62, 6.87)	(6.58, 19.25)	**(5.4, 15.25)**	(4.62, 12.26)	**(5.33, 13.44)**
Income = $25,000					
	2.37	6.13	**4.87**	3.97	**4.81**
Now 45: Sudden death at 45	(.65, 4.1)	(1.64, 10.91)	**(1.4, 8.58)**	(1.17, 6.86)	**(2.09, 7.76)**
	1.14	1.71	**1.54**	1.41	**1.82**
Now 45: at 45: 1 yr sick; recov	(.13, 2.15)	(.19, 3.27)	**(.15, 2.95)**	(0.15, 2.7)	**(.39, 3.37)**
	1.23	2.16	**1.86**	1.65	**2.29**
Now 45: at 45: 5 yrs sick; recov	(.31, 2.12)	(.63, 3.71)	**(.5, 3.25)**	(0.42, 2.9)	**(.86, 3.81)**
	3.21	7.61	**6.18**	5.13	**5.83**
Now 45: at 45: 1 yr sick; then die	(1.78, 4.89)	(3.26, 12.26)	**(2.86, 9.85)**	(2.51, 7.97)	**(3.21, 8.63)**
	3.35	9.23	**7.39**	6.01	**6.63**
Now 45: at 45: 5 yrs sick; then die	(1.91, 5.09)	(4.66, 14.14)	**(3.87, 11.23)**	(3.34, 9.05)	**(3.81, 9.88)**
Income = $67,500					
	4.51	11.9	**9.37**	7.6	**9.26**
Now 45: Sudden death at 45	(1.51, 7.51)	(4.11, 20.2)	**(3.36, 15.82)**	(2.74, 12.63)	**(4.56, 14.39)**

Now 45: at 45: 1 yr sick; recov	1.97 (.23, 3.72)	2.95 (.32, 5.65)	**2.65 (.26, 5.09)**	2.44 (0.25, 4.67)	**3.14 (.66, 5.83)**
Now 45: at 45: 5 yrs sick; recov	2.12 (.54, 3.67)	3.72 (1.08, 6.42)	**3.2 (.86, 5.62)**	2.85 (0.72, 5.02)	**3.96 (1.48, 6.59)**
Now 45: at 45: 1 yr sick; then die	5.92 (3.43, 8.84)	14.42 (6.86, 22.51)	**11.59 (5.83, 17.99)**	9.57 (5.01, 14.51)	**10.99 (6.44, 15.85)**
Now 45: at 45: 5 yrs sick; then die	6.02 (3.51, 9.05)	17.03 (9.09, 25.6)	**13.52 (7.39, 20.22)**	10.92 (6.27, 16.23)	**12.19 (7.3, 17.86)**
Latencies (income=$42K)					
Now 35: Sudden death now	1.28 (-1.21, 3.91)	-1.07 (-8.09, 5.52)	**-.29 (-5.61, 4.71)**	0.29 (-3.83, 4.38)	**.72 (-4.3, 5.92)**
Now 35: Sudden death at 40	1.92 (.14, 3.79)	-.61 (-6.6, 5)	**.31 (-3.81, 4.4)**	0.98 (-2.12, 4.02)	**1.17 (-2.79, 5.14)**
Now 35: Sudden death at 50	1.71 (.51, 3.08)	.33 (-3.8, 4.42)	**1.32 (-1.21, 3.96)**	1.86 (-0.03, 3.76)	**1.91 (-.61, 4.44)**
Now 35: Sudden death at 60	.8 (.2, 1.45)	1.2 (-1.55, 4.02)	**1.93 (.12, 3.84)**	2.02 (0.57, 3.66)	**2.3 (.32, 4.37)**
Now 35: Sudden death at 70	.25 (.06, .47)	1.87 (.01, 3.78)	**2 (.55, 3.63)**	1.56 (0.45, 2.82)	**2.19 (.43, 4)**
Now 35: Sudden death at 80	.05 (.01, .1)	1.94 (.54, 3.5)	**1.34 (.36, 2.43)**	0.72 (0.18, 1.32)	**1.4 (.21, 2.63)**
Now 35: now: 1 yr sick; recov	.8 (-.64, 2.31)	1.86 (-.49, 4.28)	**1.33 (-.74, 3.48)**	1.07 (-0.8, 3.05)	**1.92 (-.29, 4.27)**
Now 35: at 40: 1 yr sick; recov	.56 (-.49, 1.65)	1.77 (-.45, 4.05)	**1.21 (-.68, 3.16)**	0.92 (-0.71, 2.66)	**1.74 (-.3, 3.87)**
Now 35: at 50: 1 yr sick; recov	.22 (-.22, .69)	1.57 (-.37, 3.55)	**.98 (-.56, 2.55)**	0.66 (-0.53, 1.91)	**1.38 (-.26, 3.11)**
Now 35: at 60: 1 yr sick; recov	.07 (-.07, .22)	1.38 (-.25, 3.08)	**.77 (-.38, 1.95)**	0.44 (-0.33, 1.24)	**1.06 (-.17, 2.35)**
Now 35: at 70: 1 yr sick; recov	.02 (-.01, .06)	1.18 (-.06, 2.48)	**.57 (-.18, 1.35)**	0.27 (-0.14, 0.7)	**.77 (-.05, 1.62)**
Now 35: at 80: 1 yr sick; recov	.01 (0, .01)	.92 (.21, 1.69)	**.38 (.04, .74)**	0.15 (0.01, 0.3)	**.47 (.11, .86)**
Now 35: now: 5 yrs sick; recov	1.08 (-.22, 2.41)	2.63 (.26, 5.01)	**1.85 (-.19, 3.94)**	1.49 (-0.3, 3.38)	**2.68 (.5, 4.89)**

Now 35: at 40: 5 yrs sick; recov	.75 (-.13, 1.67)	2.49 (.28, 4.74)	**1.68 (-.15, 3.57)**	1.28 (-0.26, 2.9)	**2.42 (.44, 4.41)**
Now 35: at 50: 5 yrs sick; recov	.29 (-.04, .64)	2.22 (.33, 4.14)	**1.35 (-.09, 2.82)**	0.9 (-0.15, 2.02)	**1.92 (.35, 3.49)**
Now 35: at 60: 5 yrs sick; recov	.09 (-.01, .19)	1.93 (.38, 3.48)	**1.04 (.01, 2.1)**	0.58 (-0.07, 1.26)	**1.45 (.34, 2.59)**
Now 35: at 70: 5 yrs sick; recov	.03 (0, .05)	1.63 (.5, 2.76)	**.75 (.13, 1.37)**	0.34 (0.03, 0.66)	**1.01 (.33, 1.72)**
Now 35: at 80: 5 yrs sick; recov	.01 (0, .01)	1.23 (.7, 1.85)	**.48 (.24, .74)**	0.18 (0.09, 0.3)	**.6 (.31, .9)**
Now 35: now: 1 yr sick; then die	4.07 (2.2, 6.18)	3.13 (-3.23, 9.41)	**3.52 (-1.11, 8.29)**	3.78 (0.14, 7.36)	**4.47 (-.13, 9.06)**
Now 35: at 40: 1 yr sick; then die	3.18 (1.76, 4.75)	2.97 (-2.39, 8.27)	**3.25 (-.4, 6.89)**	3.39 (0.93, 5.91)	**4.07 (.55, 7.65)**
Now 35: at 50: 1 yr sick; then die	1.82 (.69, 3.18)	2.86 (-.76, 6.51)	**2.96 (.8, 5.18)**	2.89 (1.39, 4.54)	**3.55 (1.52, 5.79)**
Now 35: at 60: 1 yr sick; then die	.75 (.22, 1.35)	2.91 (.6, 5.27)	**2.76 (1.27, 4.43)**	2.37 (1.02, 3.97)	**3.13 (1.46, 5.01)**
Now 35: at 70: 1 yr sick; then die	.23 (.06, .41)	2.91 (1.42, 4.59)	**2.34 (.98, 3.97)**	1.61 (0.55, 2.84)	**2.54 (.87, 4.28)**
Now 35: at 80: 1 yr sick; then die	.05 (.01, .08)	2.34 (1.08, 3.88)	**1.35 (.46, 2.36)**	0.67 (0.19, 1.19)	**1.42 (.34, 2.53)**
Now 35: now: 5 yrs sick; then die	6.86 (4.58, 9.62)	9.27 (2.43, 16.3)	**8.83 (3.66, 14.28)**	8.38 (4.52, 12.78)	**9.66 (5.12, 14.86)**
Now 35: at 40: 5 yrs sick; then die	4.58 (3.06, 6.45)	8.54 (2.85, 14.49)	**7.75 (3.9, 12.08)**	6.97 (4.24, 10.18)	**8.49 (4.88, 12.61)**
Now 35: at 50: 5 yrs sick; then die	1.76 (.89, 2.76)	7.26 (3.42, 11.44)	**5.93 (3.76, 8.44)**	4.71 (3.17, 6.64)	**6.48 (4.46, 8.88)**
Now 35: at 60: 5 yrs sick; then die	.56 (.22, .95)	6.12 (3.76, 8.87)	**4.4 (2.93, 6.17)**	2.98 (1.73, 4.42)	**4.77 (3.22, 6.6)**
Now 35: at 70: 5 yrs sick; then die	.14 (.05, .25)	4.84 (3.29, 6.79)	**2.88 (1.66, 4.28)**	1.55 (0.73, 2.54)	**3.1 (1.67, 4.67)**
Now 35: at 80: 5 yrs sick; then die	.02 (.01, .03)	2.26 (1.36, 3.33)	**.93 (.47, 1.48)**	0.37 (0.16, 0.6)	**1.04 (.47, 1.63)**
Now 65: Sudden death now	3.53	7.25	**6.16**	5.31	**5.91**

	(.91, 6.28)	(.61, 13.59)	**(.76, 11.45)**	(0.84, 9.72)	**(1.61, 10.24)**
	1.8	5.79	**4.62**	3.69	**4.37**
Now 65: Sudden death at 70	(.23, 3.42)	(.86, 10.73)	**(.91, 8.41)**	(0.93, 6.55)	**(1.4, 7.29)**
	.2	2.92	**1.95**	1.25	**1.73**
Now 65: Sudden death at 80	(-1.07, 1.4)	(.76, 5.2)	**(.35, 3.63)**	(-0.23, 2.75)	**(-.07, 3.49)**
	-.02	.15	**.03**	-0.02	**-.05**
Now 65: Sudden death at 90	(-.22, .17)	(-1.09, 1.33)	**(-.94, .95)**	(-0.73, 0.66)	**(-1.16, 1.02)**
	2.88	3.93	**3.68**	3.47	**3.83**
Now 65: now: 1 yr sick; recov	(1.23, 4.64)	(1.52, 6.59)	**(1.49, 6.05)**	(1.44, 5.71)	**(1.7, 6.24)**
	2.03	3.54	**3.21**	2.9	**3.33**
Now 65: at 70: 1 yr sick; recov	(.87, 3.3)	(1.34, 5.92)	**(1.28, 5.33)**	(1.19, 4.76)	**(1.42, 5.44)**
	.7	2.51	**2.07**	1.67	**2.14**
Now 65: at 80: 1 yr sick; recov	(.3, 1.16)	(.86, 4.29)	**(.78, 3.5)**	(0.66, 2.79)	**(.86, 3.52)**
	2.63	3.81	**3.6**	3.38	**3.85**
Now 65: now: 5 yrs sick; recov	(1.21, 4.07)	(1.45, 6.27)	**(1.49, 5.77)**	(1.44, 5.33)	**(1.77, 6.07)**
	1.76	3.37	**3.06**	2.75	**3.28**
Now 65: at 70: 5 yrs sick; recov	(.81, 2.72)	(1.24, 5.55)	**(1.24, 4.92)**	(1.16, 4.34)	**(1.51, 5.16)**
	.49	2.06	**1.68**	1.32	**1.81**
Now 65: at 80: 5 yrs sick; recov	(.22, .77)	(.68, 3.46)	**(.61, 2.77)**	(0.52, 2.13)	**(.77, 2.9)**
	2.47	5.88	**4.87**	4.08	**3.67**
Now 65: now: 1 yr sick; then die	(-.01, 5)	(-.91, 12.53)	**(-.68, 10.24)**	(-0.46, 8.34)	**(-.45, 8.06)**
	1.3	4.65	**3.65**	2.86	**2.73**
Now 65: at 70: 1 yr sick; then die	(.05, 2.54)	(-.41, 9.47)	**(-.06, 7.2)**	(0.27, 5.57)	**(-.04, 5.45)**
	.14	2.2	**1.46**	0.93	**.98**
Now 65: at 80: 1 yr sick; then die	(-1.02, 1.2)	(.27, 4.14)	**(.16, 2.78)**	(-0.36, 2.14)	**(-.51, 2.54)**
	.72	.05	**.03**	0.02	**.15**
Now 65: now: 5 yrs sick; then die	(-1.83, 3.25)	(-.74, .79)	**(-.56, .58)**	(-0.38, 0.4)	**(-4.45, 4.52)**
	.38	3.3	**2.48**	1.86	**-.05**
Now 65: at 70: 5 yrs sick; then die	(-.91, 1.63)	(-4.6, 10.86)	**(-3.9, 8.59)**	(-3.17, 6.73)	**(-2.98, 2.9)**
	.09	2.39	**1.7**	1.2	**-.28**
Now 65: at 80: 5 yrs sick; then die	(-.63, .75)	(-3.42, 8.01)	**(-2.4, 5.68)**	(-1.65, 4.07)	**(-1.78, 1.16)**
Age Profile for specified illnesses					

Now 25: at 25: sudden death	-2.06 (-6.95, 2.57)	-19.75 (-36.15, -5.49)	**-13.28 (-25.46, -2.9)**	-8.97 (-18.1, -1.04)	**-10.54 (-22.1, .32)**
Now 30: at 30: sudden death	-.22 (-3.57, 3.15)	-9.2 (-19.01, .08)	**-6.01 (-13.54, 1.01)**	-3.82 (-9.74, 1.77)	**-4.23 (-11.74, 3.3)**
Now 35: at 35: sudden death	1.28 (-1.21, 3.91)	-1.07 (-8.09, 5.52)	**-.29 (-5.61, 4.71)**	0.29 (-3.83, 4.38)	**.72 (-4.3, 5.92)**
Now 40: at 40: sudden death	2.46 (.17, 4.78)	4.85 (-1.02, 10.95)	**3.98 (-.62, 8.84)**	3.41 (-0.35, 7.23)	**4.37 (.42, 8.38)**
Now 45: at 45: sudden death	3.31 (1, 5.61)	8.63 (2.64, 15.01)	**6.82 (2.2, 11.78)**	5.55 (1.81, 9.41)	**6.74 (3.12, 10.68)**
Now 50: at 50: sudden death	3.84 (1.55, 6.12)	10.57 (4.62, 17.14)	**8.36 (3.78, 13.49)**	6.76 (3.02, 10.88)	**7.98 (4.37, 11.92)**
Now 55: at 55: sudden death	4.05 (1.82, 6.37)	10.81 (4.98, 16.98)	**8.67 (3.99, 13.53)**	7.07 (3.28, 10.95)	**8.15 (4.44, 12.2)**
Now 60: at 60: sudden death	3.94 (1.61, 6.45)	9.61 (3.76, 15.55)	**7.88 (3.04, 12.8)**	6.55 (2.57, 10.51)	**7.39 (3.59, 11.26)**
Now 65: at 65: sudden death	3.53 (.91, 6.28)	7.25 (.61, 13.59)	**6.16 (.76, 11.45)**	5.31 (0.84, 9.72)	**5.91 (1.61, 10.24)**
Now 70: at 70: sudden death	2.87 (-.7, 6.55)	4.16 (-4.1, 12.64)	**3.8 (-3.08, 10.84)**	3.51 (-2.42, 9.44)	**4 (-1.59, 9.24)**
Now 75: at 75: sudden death	2.04 (-2.99, 6.95)	.82 (-9.85, 11.32)	**1.16 (-7.85, 10.12)**	1.44 (-6.21, 9.17)	**2.05 (-5.46, 8.98)**
Now 80: at 80: sudden death	1.09 (-5.7, 7.74)	-2.68 (-16.16, 10.91)	**-1.66 (-13.29, 9.81)**	-0.81 (-10.77, 9.16)	**.16 (-9.58, 9.16)**
Now 25: at 35: 5 yrs sick then die	6 (3.82, 8.53)	1.06 (-8.15, 9.89)	**5.27 (.48, 10.4)**	7.18 (4.07, 10.65)	**8.31 (3.36, 13.66)**
Now 30: at 40: 5 yrs sick then die	4.32 (2.77, 6.23)	5.25 (-.7, 11.6)	**6.32 (2.83, 10.32)**	6.51 (4.2, 9.23)	**8.06 (4.54, 12.11)**
Now 35: at 45: 5 yrs sick then die	2.91 (1.73, 4.3)	7.87 (3.23, 12.82)	**6.79 (3.95, 10.09)**	5.76 (3.88, 7.98)	**7.43 (4.75, 10.48)**
Now 40: at 50: 5 yrs sick then die	1.76 (.81, 2.84)	9.08 (4.77, 13.56)	**6.69 (3.99, 9.74)**	4.92 (3.15, 6.99)	**6.44 (4.25, 8.98)**
Now 45: at 55: 5 yrs sick then die	.89 (-.05, 1.86)	8.95 (4.87, 13.37)	**6.05 (3.51, 9)**	3.99 (2.36, 5.96)	**5.12 (3.21, 7.34)**
Now 50: at 60: 5 yrs sick then die	.29 (-.69, 1.26)	7.86 (4.04, 12.21)	**5.01 (2.64, 7.69)**	3.03 (1.48, 4.83)	**3.63 (1.84, 5.66)**
Now 55: at 65: 5 yrs sick then die	-.02 (-1.02, .92)	5.99 (2.29, 10.02)	**3.68 (1.22, 6.27)**	2.08 (0.52, 3.79)	**2.08 (.31, 3.9)**

	-.06	3.7	**2.24**	1.23	**.69**
Now 60: at 70: 5 yrs sick then die	(-1.09, .89)	(-.21, 7.54)	**(-.22, 4.75)**	(-0.38, 2.81)	**(-1.03, 2.44)**
	.18	1.49	**.99**	0.65	**-.22**
Now 65: at 75: 5 yrs sick then die	(-.92, 1.17)	(-2.23, 5.15)	**(-1.31, 3.33)**	(-1, 2.23)	**(-2.03, 1.58)**
	.65	.09	**.39**	0.58	**-.18**
Now 70: at 80: 5 yrs sick then die	(-.6, 1.86)	(-3.24, 3.3)	**(-1.84, 2.57)**	(-1.08, 2.28)	**(-2.16, 1.69)**
	1.24	.16	**.81**	1.19	**1.04**
Now 75: at 85: 5 yrs sick then die	(-.26, 2.74)	(-2.8, 3.1)	**(-1.51, 3.1)**	(-0.76, 3.2)	**(-1.38, 3.41)**
	1.72	1.39	**1.93**	2.15	**2.9**
Now 80: at 90: 5 yrs sick then die	(.15, 3.4)	(-1.65, 4.47)	**(-.61, 4.76)**	(-0.23, 4.69)	**(-.23, 5.96)**
		1.36	**.46**	0.15	**.56**
Now 25: ill 6 mo die 6 mo early	msg	(.71, 2.06)	**(.24, .71)**	(0.08, 0.23)	**(.3, .85)**
	.01	1	**.38**	0.14	**.44**
Now 30: ill 6 mo die 6 mo early	(0, .01)	(.49, 1.59)	**(.17, .6)**	(0.06, 0.22)	**(.2, .68)**
		.63	**.25**	0.1	**.26**
Now 35: ill 6 mo die 6 mo early	msg	(.22, 1.08)	**(.08, .44)**	(0.03, 0.17)	**(.05, .48)**
		.29	**.12**	0.05	**.08**
Now 40: ill 6 mo die 6 mo early	msg	(-.1, .7)	**(-.07, .31)**	(-0.03, 0.13)	**(-.14, .31)**
			-.03	-0.02	**-.11**
Now 45: ill 6 mo die 6 mo early	msg	msg	**(-.26, .2)**	(-0.13, 0.09)	**(-.4, .16)**
	-.01	-.22	**-.15**	-0.09	**-.27**
Now 50: ill 6 mo die 6 mo early	(-.03, .01)	(-.75, .29)	**(-.45, .14)**	(-0.25, 0.07)	**(-.64, .06)**
	-.02	-.3	**-.21**	-0.13	**-.35**
Now 55: ill 6 mo die 6 mo early	(-.05, .01)	(-.9, .27)	**(-.58, .15)**	(-0.34, 0.08)	**(-.81, .05)**
	-.03	-.23	**-.18**	-0.12	**-.29**
Now 60: ill 6 mo die 6 mo early	(-.08, .03)	(-.92, .43)	**(-.64, .27)**	(-0.41, 0.16)	**(-.84, .2)**
		.04	**.01**		**-.02**
Now 65: ill 6 mo die 6 mo early	msg	(-.78, .81)	**(-.59, .58)**	msg	**(-.69, .63)**
	.1	.55	**.41**	0.3	**.55**
Now 70: ill 6 mo die 6 mo early	(-.12, .31)	(-.52, 1.61)	**(-.43, 1.26)**	(-0.33, 0.93)	**(-.43, 1.51)**
	.33	1.31	**1.06**	0.83	**1.48**
Now 75: ill 6 mo die 6 mo early	(-.06, .76)	(-.17, 2.92)	**(-.17, 2.37)**	(-0.13, 1.86)	**(.07, 2.96)**
	.73	2.31	**1.94**	1.58	**2.74**
Now 80: ill 6 mo die 6 mo early	(.07, 1.45)	(.24, 4.55)	**(.19, 3.83)**	(0.14, 3.12)	**(.68, 4.92)**
	.07	1.45	**.49**	0.17	**1.28**
Now 25: at 25: 1 year sick; recover	(-1.7, 1.89)	(-1.39, 4.46)	**(-2.08, 3.12)**	(-2.14, 2.54)	**(-1.45, 4.14)**

Now 30: at 30: 1 year sick; recover	.44 (-1.17, 2.07)	1.68 (-.93, 4.4)	**.93 (-1.35, 3.29)**	0.64 (-1.45, 2.75)	**1.63 (-.82, 4.2)**
Now 35: at 35: 1 year sick; recover	.8 (-.64, 2.31)	1.86 (-.49, 4.28)	**1.33 (-.74, 3.48)**	1.07 (-0.8, 3.05)	**1.92 (-.29, 4.27)**
Now 40: at 40: 1 year sick; recover	1.16 (-.2, 2.56)	2.06 (-.09, 4.26)	**1.7 (-.21, 3.69)**	1.49 (-0.24, 3.31)	**2.19 (.22, 4.37)**
Now 45: at 45: 1 year sick; recover	1.51 (.18, 2.86)	2.27 (.25, 4.36)	**2.05 (.2, 3.92)**	1.88 (0.19, 3.59)	**2.42 (.51, 4.49)**
Now 50: at 50: 1 year sick; recover	1.86 (.52, 3.2)	2.53 (.48, 4.68)	**2.39 (.53, 4.25)**	2.25 (0.52, 3.96)	**2.66 (.76, 4.67)**
Now 55: at 55: 1 year sick; recover	2.2 (.79, 3.65)	2.88 (.68, 5.06)	**2.75 (.78, 4.74)**	2.63 (0.83, 4.46)	**2.94 (.96, 5.01)**
Now 60: at 60: 1 year sick; recover	2.54 (1.01, 4.16)	3.33 (1.03, 5.76)	**3.17 (1.07, 5.4)**	3.02 (1.09, 5.06)	**3.31 (1.27, 5.52)**
Now 65: at 65: 1 year sick; recover	2.88 (1.23, 4.64)	3.93 (1.52, 6.59)	**3.68 (1.49, 6.05)**	3.47 (1.44, 5.71)	**3.83 (1.7, 6.24)**
Now 70: at 70: 1 year sick; recover	3.26 (1.48, 5.15)	4.68 (1.98, 7.47)	**4.33 (1.88, 6.94)**	4.03 (1.78, 6.41)	**4.57 (2.22, 7.19)**
Now 75: at 75: 1 year sick; recover	3.73 (1.76, 5.79)	5.64 (2.62, 8.85)	**5.19 (2.44, 8.09)**	4.77 (2.21, 7.51)	**5.62 (3.05, 8.68)**
Now 80: at 80: 1 year sick; recover	4.34 (2.12, 6.66)	6.85 (3.41, 10.59)	**6.27 (3.13, 9.67)**	5.72 (2.93, 8.91)	**7.00 (3.99, 10.56)**
Income effects:					
Now 45: Sudden death at 45 for income $10K	1.34 (.31, 2.37)	3.42 (.72, 6.28)	**2.73 (.65, 4.95)**	2.24 (0.55, 3.97)	**$ 2.70 (1.07, 4.46)**
Now 45: Sudden death at 45 for income $20K	2.06 (.54, 3.58)	5.31 (1.34, 9.53)	**4.22 (1.16, 7.5)**	3.44 (0.97, 6)	**4.17 (1.77, 6.77)**
Now 45: Sudden death at 45 for income $30K	2.66 (.75, 4.57)	6.91 (1.94, 12.2)	**5.48 (1.64, 9.59)**	4.46 (1.36, 7.66)	**5.41 (2.41, 8.68)**
Now 45: Sudden death at 45 for income $40K	3.2 (.96, 5.44)	8.35 (2.53, 14.56)	**6.61 (2.11, 11.43)**	5.37 (1.74, 9.14)	**6.53 (3.01, 10.36)**
Now 45: Sudden death at 45 for income $50K	3.7 (1.17, 6.24)	9.7 (3.11, 16.73)	**7.66 (2.57, 13.12)**	6.22 (2.11, 10.48)	**7.57 (3.58, 11.91)**
Now 45: Sudden death at 45 for income $60K	4.17 (1.37, 6.98)	10.98 (3.68, 18.75)	**8.65 (3.02, 14.7)**	7.02 (2.47, 11.74)	**8.55 (4.15, 13.36)**
Now 45: Sudden death at 45 for income $70K	4.62 (1.56, 7.68)	12.2 (4.26, 20.67)	**9.6 (3.47, 16.19)**	7.79 (2.83, 12.92)	**9.49 (4.69, 14.73)**

Now 45: Sudden death at 45 for income $80K	5.05 (1.76, 8.34)	13.38 (4.83, 22.5)	**10.52 (3.92, 17.61)**	8.52 (3.19, 14.05)	**10.4 (5.23, 16.04)**
Now 45: Sudden death at 45 for income $90K	5.47 (1.95, 8.98)	14.52 (5.39, 24.26)	**11.4 (4.36, 18.97)**	9.23 (3.54, 15.14)	**11.28 (5.76, 17.3)**
Now 45: Sudden death at 45 for income $100K	5.87 (2.15, 9.6)	15.63 (5.96, 25.95)	**12.26 (4.8, 20.29)**	9.92 (3.89, 16.18)	**12.13 (6.28, 18.51)**
Now 45: Sudden death at 45 for income $110K	6.26 (2.34, 10.19)	16.72 (6.52, 27.6)	**13.1 (5.23, 21.57)**	10.6 (4.23, 17.2)	**12.96 (6.8, 19.69)**
Now 45: Sudden death at 45 for income $120K	6.65 (2.53, 10.77)	17.78 (7.08, 29.2)	**13.93 (5.67, 22.81)**	11.26 (4.58, 18.18)	**13.78 (7.31, 20.84)**
Now 45: Sudden death at 45 for income $130K	7.02 (2.72, 11.33)	18.82 (7.64, 30.76)	**14.73 (6.1, 24.02)**	11.9 (4.92, 19.14)	**14.58 (7.82, 21.96)**
Now 45: Sudden death at 45 for income $140K	7.39 (2.9, 11.88)	19.85 (8.2, 32.29)	**15.52 (6.53, 25.2)**	12.53 (5.26, 20.08)	**15.36 (8.32, 23.05)**
Now 45: Sudden death at 45 for income $150K	7.75 (3.09, 12.42)	20.86 (8.76, 33.78)	**16.3 (6.96, 26.35)**	13.16 (5.6, 20.99)	**16.14 (8.82, 24.12)**
Now 45: Sudden death at 45 for income $160K	8.11 (3.28, 12.94)	21.86 (9.32, 35.25)	**17.07 (7.39, 27.49)**	13.77 (5.94, 21.89)	**16.9 (9.32, 25.18)**

Table 5-13 Linear versus Box-Cox transformation of net income
(See Section . Three types of exclusion criteria, limited scenario adjustment corrections.)

	Income retained during sick-years:	**1 Linear in net income**	2 Box-Cox in net income
	Constructed Variable		$\lambda = 0.45$
	Basic variables:		
β_0	$\left(Y_i - c_i^j\right)^{(0.45)} cterm_i^j - \left(Y_i\right)^{(0.45)} yterm_i^j$	.00003914 (8.43)***	.01141 (9.86)***
α_{10}	$\Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$	-60.78 (5.45)***	-58.35 (5.28)***
α_{20}	$\Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	66.51 (1.98)**	67.25 (2.00)**
α_{21}	… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	-2.14 (3.00)***	-2.159 (3.02)***
α_{30}	$\Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-536.4 (2.92)***	-549.4 (2.99)***
α_{31}	… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	19.59 (2.64)***	20.38 (2.74)***
α_{32}	… $age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-.1823 (2.57)**	-.1917 (2.70)***
α_{40}	$\Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	168.2 (1.98)**	176.1 (2.07)**
α_{41}	… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	-7.176 (2.08)**	-7.497 (2.17)**
α_{42}	… $age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	.07139 (2.13)**	.07512 (2.24)**
α_{50}	$\Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	114 (1.52)	113.3 (1.51)
α_{51}	… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	-4.272 (1.46)	-4.316 (1.47)
α_{52}	… $age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	.04827 (1.75)*	.04902 (1.77)*
	Status quo effect variable:		
δ	1(neither program)	-.2375	-.234

		(5.56)***	(5.52)***
	Systematic selection correction term:		
α_{13}	$\left[P(sel_i)-\overline{P}\right]\times\Delta\Pi_i^{jS}\left[\log\left(pdvi_i^j+1\right)\right]$	3.284 (2.22)**	3.294 (2.22)**
	Scenario adjustment variables[b]		
	$age_{i0}\times\Delta\Pi_i^{jS}\log\left(pdvi_i^j+1\right)\times benefit\_never_i^j$	22.63 (5.52)***	22.36 (5.46)***
	$age_{i0}^2\times\Delta\Pi_i^{jS}\log\left(pdvi_i^j+1\right)\times benefit\_never_i^j$	-.2738 (4.62)***	-.2702 (4.56)***
	$age_{i0}^2\times\Delta\Pi_i^{jS}\log\left(pdvl_i^j+1\right)\times benefit\_never_i^j$	.1326 (3.98)***	.1338 (4.03)***
	$\Delta\Pi_i^{jS}\left[\log\left(pdvl_i^j+1\right)\right]^2\times benefit\_never_i^j$	708.7 (6.45)***	707.1 (6.46)***
	$age_{i0}\times\Delta\Pi_i^{jS}\left[\log\left(pdvl_i^j+1\right)\right]^2\times benefit\_never_i^j$	-13.2 (5.71)***	-13.21 (5.73)***
	$age_{i0}\times\Delta\Pi_i^{jS}\left[\log\left(pdvi_i^j+1\right)\right]\times\left[\log\left(pdvl_i^j+1\right)\right]\times benefit\_never_i^j$	-26.74 (5.30)***	-26.54 (5.27)***
	$age_{i0}^2\times\Delta\Pi_i^{jS}\left[\log\left(pdvi_i^j+1\right)\right]\times\left[\log\left(pdvl_i^j+1\right)\right]\times benefit\_never_i^j$	.3694 (4.71)***	.3662 (4.69)***
	$age_{i0}^2\times\Delta\Pi_i^{jS}\log\left(pdvi_i^j+1\right)\times\log\left(\left\|LEdiff<0\right\|+1\right)$	.003364 (2.47)**	.003011 (2.20)**
	$age_{i0}\times\Delta\Pi_i^{jS}\log\left(pdvr_i^j+1\right)\times\log\left(\left\|LEdiff<0\right\|+1\right)$	.3644 (2.22)**	.3672 (2.23)**
	Max LogL	-11486.923	-11471.366
	Alternatives	22,560	22,560

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

[a] For the linear model, the implicit Box-Cox parameter is constrained to be 1. When we determine the best alternative value for this parameter by a line-search, according to the maximized value of the log-likelihood, we settle on a value of 0.45. At this value of the parameter, the log likelihood improves by 15.56 points, so between the two values, the Box-Cox specification is preferred for this sample.

Table 5-14 WTP with linear versus Box-Cox transformations of net income
(See Section. Based on parameter estimates from the models in Table 5-13.)

Illness profile: age 45 now; …at 45:	**linear**	Box-Cox $\lambda = 0.45$
Income = $42,000		
Now 45: Sudden death at 45	$ 4.89 (3.27, 6.93)	$ 5.96 (3.99, 7.96)
Now 45: at 45: 1 yr sick; recov	3.34 (1.58, 5.14)	3.91 (1.77, 5.99)
Now 45: at 45: 5 yrs sick; recov	4.69 (2.92, 6.6)	5.43 (3.44, 7.49)
Now 45: at 45: 1 yr sick; then die	4.81 (3.28, 6.66)	5.85 (4.04, 7.86)
Now 45: at 45: 5 yrs sick; then die	4.51 (2.78, 6.5)	5.41 (3.48, 7.45)
Income= $25,000		
Now 45: Sudden death at 45	4.60 (2.97, 6.63)	4.22 (2.75, 5.72)
Now 45: at 45: 1 yr sick; recov	3.35 (1.58, 5.15)	2.93 (1.33, 4.5)
Now 45: at 45: 5 yrs sick; recov	4.69 (2.92, 6.6)	4.07 (2.59, 5.62)
Now 45: at 45: 1 yr sick; then die	4.53 (3, 6.38)	4.15 (2.79, 5.66)
Now 45: at 45: 5 yrs sick; then die	4.29 (2.56, 6.28)	3.87 (2.42, 5.4)
Income= $67,500		
Now 45: Sudden death at 45	5.34 (3.71, 7.37)	8.25 (5.68, 10.85)
Now 45: at 45: 1 yr sick; recov	3.34 (1.57, 5.14)	5.07 (2.3, 7.79)
Now 45: at 45: 5 yrs sick; recov	4.69 (2.92, 6.59)	7.05 (4.47, 9.73)
Now 45: at 45: 1 yr sick; then die	5.23 (3.7, 7.08)	8.07 (5.71, 10.68)
Now 45: at 45: 5 yrs sick; then die	4.84 (3.11, 6.83)	7.4 (4.89, 10.06)
Latency (income=$42K)		
Now 35: Sudden death now	5.54 (3.42, 7.89)	6.72 (4.2, 9.12)
Now 35: Sudden death at 40	5.04	6.03

	(3.37, 6.92)	(3.98, 8.03)
Now 35: Sudden death at 50	4.1 (2.91, 5.43)	4.81 (3.42, 6.25)
Now 35: Sudden death at 60	3.17 (2.25, 4.18)	3.68 (2.6, 4.86)
Now 35: Sudden death at 70	2.19 (1.44, 2.97)	2.53 (1.62, 3.49)
Now 35: Sudden death at 80	1.08 (.6, 1.56)	1.25 (.67, 1.86)
Now 35: now: 1 yr sick; recov	1.83 (-.13, 3.8)	2.1 (-.23, 4.46)
Now 35: at 40: 1 yr sick; recov	1.6 (-.17, 3.39)	1.84 (-.31, 4.01)
Now 35: at 50: 1 yr sick; recov	1.2 (-.24, 2.63)	1.39 (-.33, 3.12)
Now 35: at 60: 1 yr sick; recov	.88 (-.2, 1.96)	1.02 (-.27, 2.31)
Now 35: at 70: 1 yr sick; recov	.62 (-.08, 1.32)	.72 (-.11, 1.57)
Now 35: at 80: 1 yr sick; recov	.39 (.11, .68)	.45 (.12, .8)
Now 35: now: 5 yrs sick; recov	3.31 (1.42, 5.23)	3.79 (1.64, 6.04)
Now 35: at 40: 5 yrs sick; recov	2.96 (1.25, 4.69)	3.39 (1.46, 5.42)
Now 35: at 50: 5 yrs sick; recov	2.29 (.97, 3.64)	2.63 (1.13, 4.2)
Now 35: at 60: 5 yrs sick; recov	1.7 (.76, 2.66)	1.96 (.87, 3.08)
Now 35: at 70: 5 yrs sick; recov	1.19 (.64, 1.79)	1.37 (.74, 2.04)
Now 35: at 80: 5 yrs sick; recov	.76 (.55, .99)	.87 (.62, 1.13)
Now 35: now: 1 yr sick; then die	5.26 (3.46, 7.29)	6.39 (4.27, 8.61)
Now 35: at 40: 1 yr sick; then die	4.89 (3.51, 6.63)	5.84 (4.13, 7.61)
Now 35: at 50: 1 yr sick; then die	4.09 (3.11, 5.28)	4.79 (3.62, 6.04)
Now 35: at 60: 1 yr sick; then die	3.21 (2.38, 4.14)	3.72 (2.73, 4.77)
Now 35: at 70: 1 yr sick; then die	2.23 (1.5, 2.97)	2.57 (1.7, 3.48)
Now 35: at 80: 1 yr sick; then die	1.08 (.64, 1.52)	1.25 (.72, 1.78)
Now 35: now: 5 yrs sick; then die	4.81	5.8

	(2.86, 6.91)	(3.55, 8.32)
Now 35: at 40: 5 yrs sick; then die	4.6 (3.08, 6.33)	5.46 (3.73, 7.42)
Now 35: at 50: 5 yrs sick; then die	4.02 (3.06, 5.18)	4.69 (3.54, 5.88)
Now 35: at 60: 5 yrs sick; then die	3.24 (2.43, 4.12)	3.75 (2.82, 4.71)
Now 35: at 70: 5 yrs sick; then die	2.25 (1.61, 2.95)	2.59 (1.83, 3.35)
Now 35: at 80: 5 yrs sick; then die	.98 (.71, 1.29)	1.13 (.8, 1.47)
Now 65: Sudden death now	2.27 (.75, 3.89)	2.79 (.79, 4.72)
Now 65: Sudden death at 70	1.97 (.82, 3.11)	2.38 (.91, 3.78)
Now 65: Sudden death at 80	1.35 (.51, 2.19)	1.61 (.65, 2.61)
Now 65: Sudden death at 90	.38 (-.11, .86)	.47 (-.1, 1.03)
Now 65: now: 1 yr sick; recov	6.19 (3.8, 8.84)	7.29 (4.53, 10.29)
Now 65: at 70: 1 yr sick; recov	5.39 (3.3, 7.72)	6.37 (3.94, 8.99)
Now 65: at 80: 1 yr sick; recov	3.57 (2.2, 5.08)	4.24 (2.64, 5.92)
Now 65: now: 5 yrs sick; recov	7 (4.7, 9.52)	8.16 (5.58, 10.91)
Now 65: at 70: 5 yrs sick; recov	6.03 (4.09, 8.18)	7.05 (4.85, 9.39)
Now 65: at 80: 5 yrs sick; recov	3.7 (2.65, 4.91)	4.33 (3.11, 5.61)
Now 65: now: 1 yr sick; then die	1.36 (-.11, 2.8)	1.67 (-.23, 3.43)
Now 65: at 70: 1 yr sick; then die	1.4 (.44, 2.37)	1.66 (.42, 2.88)
Now 65: at 80: 1 yr sick; then die	1.24 (.5, 1.95)	1.47 (.64, 2.35)
Now 65: at 90: 1 yr sick; then die	.53 (.24, .81)	.62 (.29, .97)
Now 65: now: 5 yrs sick; then die	.32 (-1.42, 2.05)	.32 (-1.76, 2.34)
Now 65: at 70: 5 yrs sick; then die	.76 (-.43, 1.93)	.84 (-.65, 2.16)
Now 65: at 80: 5 yrs sick; then die	1.35 (.57, 2.08)	1.55 (.66, 2.41)

Age profiles

Now 25: at 25: sudden death	5.52 (1.54, 9.81)	6.64 (1.94, 11.46)
Now 30: at 30: sudden death	5.61 (2.75, 8.63)	6.78 (3.36, 10.14)
Now 35: at 35: sudden death	5.54 (3.42, 7.89)	6.72 (4.2, 9.12)
Now 40: at 40: sudden death	5.3 (3.53, 7.35)	6.44 (4.33, 8.49)
Now 45: at 45: sudden death	4.89 (3.27, 6.93)	5.96 (3.99, 7.96)
Now 50: at 50: sudden death	4.36 (2.79, 6.24)	5.33 (3.26, 7.26)
Now 55: at 55: sudden death	3.72 (2.2, 5.45)	4.56 (2.66, 6.47)
Now 60: at 60: sudden death	3.01 (1.5, 4.62)	3.69 (1.82, 5.53)
Now 65: at 65: sudden death	2.27 (.75, 3.89)	2.79 (.79, 4.72)
Now 70: at 70: sudden death	1.59 (-.21, 3.38)	1.97 (-.42, 4.28)
Now 75: at 75: sudden death	1.06 (-1.41, 3.32)	1.32 (-1.82, 4.36)
Now 80: at 80: sudden death	.66 (-2.6, 3.74)	.86 (-3.23, 4.82)
Now 25: at 35: 5 yrs sick then die	4.5 (2.34, 6.96)	5.31 (2.74, 8.1)
Now 30: at 40: 5 yrs sick then die	4.51 (2.94, 6.27)	5.3 (3.49, 7.31)
Now 35: at 45: 5 yrs sick then die	4.33 (3.16, 5.73)	5.09 (3.72, 6.6)
Now 40: at 50: 5 yrs sick then die	3.98 (2.98, 5.22)	4.67 (3.56, 5.93)
Now 45: at 55: 5 yrs sick then die	3.48 (2.59, 4.61)	4.06 (3.02, 5.22)
Now 50: at 60: 5 yrs sick then die	2.88 (2.02, 3.9)	3.34 (2.36, 4.41)
Now 55: at 65: 5 yrs sick then die	2.22 (1.41, 3.13)	2.55 (1.6, 3.52)
Now 60: at 70: 5 yrs sick then die	1.59 (.79, 2.47)	1.81 (.86, 2.73)
Now 65: at 75: 5 yrs sick then die	1.11 (.23, 1.99)	1.25 (.18, 2.26)
Now 70: at 80: 5 yrs sick then die	.97 (-.01, 1.94)	1.13 (-.06, 2.29)
Now 75: at 85: 5 yrs sick then die	1.25 (.07, 2.39)	1.5 (.09, 2.85)
Now 80: at 90: 5 yrs sick then die	1.77	2.15

	(.42, 3.16)	(.6, 3.68)
Now 25: ill 6 mo die 6 mo early	.25 (.14, .37)	.3 (.17, .44)
Now 30: ill 6 mo die 6 mo early	.24 (.14, .34)	.28 (.17, .4)
Now 35: ill 6 mo die 6 mo early	.21 (.12, .29)	.24 (.14, .34)
Now 40: ill 6 mo die 6 mo early	.19 (.1, .29)	.22 (.1, .32)
Now 45: ill 6 mo die 6 mo early	.18 (.06, .29)	.2 (.05, .34)
Now 50: ill 6 mo die 6 mo early	.17 (.01, .32)	.19 (0, .37)
Now 55: ill 6 mo die 6 mo early	.19 (0, .38)	.21 (-.01, .44)
Now 60: ill 6 mo die 6 mo early	.26 (.03, .48)	.3 (.03, .56)
Now 65: ill 6 mo die 6 mo early	.39 (.1, .68)	.46 (.13, .81)
Now 70: ill 6 mo die 6 mo early	.61 (.23, 1)	.74 (.3, 1.2)
Now 75: ill 6 mo die 6 mo early	.91 (.38, 1.46)	1.13 (.5, 1.76)
Now 80: ill 6 mo die 6 mo early	1.29 (.53, 2.07)	1.61 (.73, 2.47)
Income effects		
Now 45: Sudden death at 45 for income $10K	4.34 (2.71, 6.37)	2.35 (1.46, 3.25)
Now 45: Sudden death at 45 for income $20K	4.51 (2.88, 6.54)	3.65 (2.35, 4.98)
Now 45: Sudden death at 45 for income $30K	4.68 (3.06, 6.72)	4.76 (3.13, 6.42)
Now 45: Sudden death at 45 for income $40K	4.86 (3.23, 6.89)	5.77 (3.85, 7.72)
Now 45: Sudden death at 45 for income $50K	5.03 (3.41, 7.07)	6.71 (4.54, 8.91)
Now 45: Sudden death at 45 for income $60K	5.21 (3.58, 7.24)	7.6 (5.2, 10.04)
Now 45: Sudden death at 45 for income $70K	5.38 (3.75, 7.41)	8.46 (5.84, 11.11)
Now 45: Sudden death at 45 for income $80K	5.55 (3.93, 7.59)	9.28 (6.47, 12.14)
Now 45: Sudden death at 45 for income $90K	5.73 (4.1, 7.76)	10.08 (7.08, 13.14)
Now 45: Sudden death at 45 for income $100K	5.9 (4.28, 7.94)	10.87 (7.68, 14.1)
Now 45: Sudden death at 45 for income $110K	6.08 (4.45, 8.11)	11.63 (8.28, 15.04)

Now 45: Sudden death at 45 for income $120K	6.25 (4.63, 8.28)	12.38 (8.86, 15.96)
Now 45: Sudden death at 45 for income $130K	6.43 (4.8, 8.46)	13.12 (9.44, 16.86)
Now 45: Sudden death at 45 for income $140K	6.6 (4.97, 8.63)	13.84 (10.01, 17.74)
Now 45: Sudden death at 45 for income $150K	6.77 (5.15, 8.81)	14.55 (10.57, 18.6)
Now 45: Sudden death at 45 for income $160K	6.95 (5.32, 8.98)	15.26 (11.13, 19.45)

[b] These estimates are based on an arbitrarily specified initial risk of 0.040 and a risk reduction to 0.036. The risk reduction is thus of magnitude 0.004 and this average WTP for a microrisk reduction is calculated across these 4000 microrisks. Results are minimally different for other baseline and reduced risks. See Table 5-17. The slight curvature of the utility function means that the average WTP amounts will differ somewhat with the size of the risk reduction over which they are calculated. Also, negative point estimates of WTP are not precluded by the formulas used to calculate them. Extreme draws from the estimated joint distribution of the utility parameters can produce negative simulated values. Since there was no opportunity for anyone to express a negative willingness to pay for any of the risk-reduction programs, we adopt a Tobit-like interpretation of the fitted WTP values and interpret negative fitted values as zero.

Table 5-15 Varying assumptions about income while sick
(See Section 5.6.10. Three sets of exclusion criteria, limited set of scenario adjustment controls.)

	Income retained during sick-years:	**1 All**	2 Half	3 None
	Constructed Variable	$\gamma_1 = 1$	$\gamma_1 = 0.5$	$\gamma_1 = 0$
	Basic variables:			
β_0	$\left(Y_i - c_i^j\right)^{(0.45)} cterm_i^j - \left(Y_i\right)^{(0.45)} yterm_i^j$	.01141 (9.86)***	.01139 (9.89)***	.01127 (9.96)***
α_{10}	$\Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$	-58.35 (5.28)***	-55.25 (5.02)***	-46.79 (4.28)***
α_{20}	$\Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	67.25 (2.00)**	66.85 (1.99)**	65.71 (1.95)*
α_{21}	… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	-2.159 (3.02)***	-2.159 (3.02)***	-2.157 (3.02)***
α_{30}	$\Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-549.4 (2.99)***	-549.2 (2.99)***	-549.1 (2.99)***
α_{31}	… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	20.38 (2.74)***	20.35 (2.74)***	20.29 (2.73)***
α_{32}	… $age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-.1917 (2.70)***	-.1915 (2.69)***	-.1907 (2.68)***
α_{40}	$\Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	176.1 (2.07)**	176.3 (2.07)**	176.8 (2.08)**
α_{41}	… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	-7.497 (2.17)**	-7.498 (2.17)**	-7.504 (2.17)**
α_{42}	… $age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	.07512 (2.24)**	.07516 (2.24)**	.07524 (2.24)**
α_{50}	$\Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	113.4 (1.51)	111.9 (1.49)	108.1 (1.44)
α_{51}	… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	-4.316 (1.47)	-4.263 (1.46)	-4.12 (1.41)
α_{52}	… $age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	.04902 (1.77)*	.04848 (1.75)*	.04701 (1.70)*
	Status quo effect variable:			
δ	1(neither program)	-.2340 (5.52)***	-.2351 (5.54)***	-.2377 (5.60)***
	Systematic selection correction term:			
α_{13}	$\left[P(sel_i) - \overline{P}\right] \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right]$	3.29	3.288	3.285

	(2.22)**	(2.22)**	(2.22)**
Scenario adjustment variables[b]			
$age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$ $\times benefit\_never_i^j$	22.36 (5.46)***	22.34 (5.46)***	22.28 (5.44)***
$age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$ $\times benefit\_never_i^j$	-.2702 (4.56)***	-.27 (4.55)***	-.2692 (4.54)***
$age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$ $\times benefit\_never_i^j$	.1338 (4.03)***	.1339 (4.03)***	.1343 (4.04)***
$\Delta\Pi_i^{jS}\left[\log\left(pdvl_i^j + 1\right)\right]^2$ $\times benefit\_never_i^j$	707.1 (6.46)***	707.2 (6.47)***	707.4 (6.47)***
$age_{i0} \times \Delta\Pi_i^{jS}\left[\log\left(pdvl_i^j + 1\right)\right]^2$ $\times benefit\_never_i^j$	-13.21 (5.73)***	-13.22 (5.74)***	-13.23 (5.74)***
$age_{i0} \times \Delta\Pi_i^{jS}\left[\log\left(pdvi_i^j + 1\right)\right]$ $\times\left[\log\left(pdvl_i^j + 1\right)\right] \times benefit\_never_i^j$	-26.54 (5.27)***	-26.51 (5.26)***	-26.45 (5.26)***
$age_{i0}^2 \times \Delta\Pi_i^{jS}\left[\log\left(pdvi_i^j + 1\right)\right]$ $\times\left[\log\left(pdvl_i^j + 1\right)\right] \times benefit\_never_i^j$	.3662 (4.69)***	.3658 (4.69)***	.3648 (4.68)***
$age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$ $\times \log\left(\left\|LEdiff < 0\right\| + 1\right)$	.00301 (2.20)**	.002985 (2.18)**	.002922 (2.14)**
$age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$ $\times \log\left(\left\|LEdiff < 0\right\| + 1\right)$	.3672 (2.23)**	.3672 (2.23)**	.3668 (2.23)**
Max LogL	-11471.372	-11471.018	-11470.341
Alternatives	22,560	**22,560**	22,560

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

[a] For comparability, we constrain the set of "basic" variables to be the same in each model as in the model for the 0.05 discount rate. However, we allow the scenario adjustment variables to be whatever the data dictate under the imposed discount rates in question, and there are minor differences across these three models.

Table 5-16 WTP with differ proportions of income while sick
(See Section 5.6.10. Based on parameter estimates from the models in Table 5-15.)

Illness profile: age 45 now; …at 45:	$\gamma_1 = 1$	$\gamma_1 = 0.5$	$\gamma_1 = 0$	Constant net income
Income = $42,000				
Now 45: Sudden death at 45	**$ 6.74 (3.12, 10.68)**	$ 6.73 (3.13, 10.65)	$ 6.73 (3.14, 10.49)	$ 6.73 (3.05, 10.76)
Now 45: at 45: 1 yr sick; recov	**2.42 (.51, 4.49)**	2.41 (.51, 4.47)	2.38 (0.52, 4.44)	2.43 (.53, 4.53)
Now 45: at 45: 5 yrs sick; recov	**3.05 (1.15, 5.07)**	3.03 (1.13, 5.02)	2.97 (1.07, 4.98)	3.08 (1.15, 5.19)
Now 45: at 45: 1 yr sick; then die	**8.09 (4.6, 11.82)**	8.03 (4.59, 11.73)	7.91 (4.52, 11.38)	8.09 (4.61, 12.01)
Now 45: at 45: 5 yrs sick; then die	**9.09 (5.33, 13.44)**	9.01 (5.27, 13.24)	8.86 (5.02, 12.82)	9.16 (5.42, 13.22)
Income= $25,000				
Now 45: Sudden death at 45	**4.81 (2.09, 7.76)**	4.8 (2.1, 7.73)	4.8 (2.1, 7.62)	5.06 (2.3, 8.08)
Now 45: at 45: 1 yr sick; recov	**1.82 (.39, 3.37)**	1.81 (.38, 3.35)	1.78 (0.37, 3.32)	1.83 (.4, 3.41)
Now 45: at 45: 5 yrs sick; recov	**2.29 (.86, 3.81)**	2.26 (.84, 3.75)	2.17 (0.74, 3.68)	2.32 (.87, 3.9)
Now 45: at 45: 1 yr sick; then die	**5.83 (3.21, 8.63)**	5.78 (3.2, 8.55)	5.68 (3.14, 8.28)	6.08 (3.48, 9.01)
Now 45: at 45: 5 yrs sick; then die	**6.63 (3.81, 9.88)**	6.55 (3.75, 9.72)	6.39 (3.52, 9.36)	6.88 (4.08, 9.91)
Income= $67,500				
Now 45: Sudden death at 45	**9.26 (4.56, 14.39)**	9.25 (4.57, 14.35)	9.25 (4.58, 14.14)	8.72 (3.93, 13.97)
Now 45: at 45: 1 yr sick; recov	**3.14 (.66, 5.83)**	3.13 (.67, 5.81)	3.12 (0.69, 5.79)	3.15 (.68, 5.88)
Now 45: at 45: 5 yrs sick; recov	**3.96 (1.48, 6.59)**	3.96 (1.5, 6.56)	3.98 (1.51, 6.6)	4 (1.48, 6.74)
Now 45: at 45: 1 yr sick; then die	**10.99 (6.44, 15.85)**	10.92 (6.44, 15.74)	10.78 (6.38, 15.3)	10.49 (5.96, 15.6)
Now 45: at 45: 5 yrs sick; then die	**12.19 (7.3, 17.86)**	12.13 (7.25, 17.63)	12.01 (7.02, 17.18)	11.89 (7.02, 17.19)
Latency (income=$42K)				
Now 35: Sudden death now	**.72 (-4.3, 5.92)**	.73 (-4.27, 5.95)	0.78 (-4.09, 5.91)	.66 (-4.52, 5.76)
Now 35: Sudden death at 40	**1.17 (-2.79, 5.14)**	1.18 (-2.74, 5.13)	1.23 (-2.66, 5.28)	1.25 (-2.69, 5.27)
Now 35: Sudden death at 50	**1.91 (-.61, 4.44)**	1.92 (-.6, 4.46)	1.96 (-0.52, 4.62)	2.07 (-.49, 4.72)

Now 35: Sudden death at 60	**2.3 (.32, 4.37)**	2.3 (.34, 4.37)	2.34 (0.36, 4.39)	2.41 (.41, 4.48)
Now 35: Sudden death at 70	**2.19 (.43, 4)**	2.2 (.43, 4.01)	2.22 (0.46, 3.96)	2.22 (.44, 4.06)
Now 35: Sudden death at 80	**1.4 (.21, 2.63)**	1.4 (.22, 2.64)	1.42 (0.23, 2.57)	1.37 (.18, 2.62)
Now 35: now: 1 yr sick; recov	**1.92 (-.29, 4.27)**	1.91 (-.3, 4.26)	1.89 (-0.35, 4.21)	1.94 (-.29, 4.32)
Now 35: at 40: 1 yr sick; recov	**1.74 (-.3, 3.87)**	1.72 (-.31, 3.86)	1.71 (-0.32, 3.82)	1.75 (-.29, 3.9)
Now 35: at 50: 1 yr sick; recov	**1.38 (-.26, 3.11)**	1.38 (-.26, 3.09)	1.38 (-0.26, 3.07)	1.39 (-.25, 3.12)
Now 35: at 60: 1 yr sick; recov	**1.06 (-.17, 2.35)**	1.06 (-.17, 2.36)	1.07 (-0.15, 2.34)	1.07 (-.16, 2.36)
Now 35: at 70: 1 yr sick; recov	**.77 (-.05, 1.62)**	.76 (-.05, 1.61)	0.77 (-0.04, 1.61)	.76 (-.05, 1.61)
Now 35: at 80: 1 yr sick; recov	**.47 (.11, .86)**	.47 (.11, .85)	0.46 (0.1, 0.85)	.46 (.09, .85)
Now 35: now: 5 yrs sick; recov	**2.68 (.5, 4.89)**	2.65 (.48, 4.87)	2.59 (0.42, 4.94)	2.71 (.5, 5.05)
Now 35: at 40: 5 yrs sick; recov	**2.42 (.44, 4.41)**	2.39 (.41, 4.39)	2.31 (0.34, 4.41)	2.45 (.45, 4.58)
Now 35: at 50: 5 yrs sick; recov	**1.92 (.35, 3.49)**	1.88 (.31, 3.43)	1.79 (0.24, 3.44)	1.94 (.37, 3.58)
Now 35: at 60: 5 yrs sick; recov	**1.45 (.34, 2.59)**	1.42 (.3, 2.53)	1.33 (0.21, 2.52)	1.46 (.34, 2.67)
Now 35: at 70: 5 yrs sick; recov	**1.01 (.33, 1.72)**	.99 (.3, 1.69)	0.92 (0.23, 1.63)	1.02 (.34, 1.73)
Now 35: at 80: 5 yrs sick; recov	**.6 (.31, .9)**	.58 (.29, .87)	0.51 (0.23, 0.81)	.6 (.32, .91)
Now 35: now: 1 yr sick; then die	**4.47 (-.13, 9.06)**	4.42 (-.2, 9.01)	4.32 (-0.23, 8.98)	4.43 (-.07, 9.04)
Now 35: at 40: 1 yr sick; then die	**4.07 (.55, 7.65)**	4.03 (.51, 7.63)	3.95 (0.42, 7.53)	4.15 (.67, 7.69)
Now 35: at 50: 1 yr sick; then die	**3.55 (1.52, 5.79)**	3.52 (1.51, 5.76)	3.48 (1.46, 5.71)	3.7 (1.68, 5.95)
Now 35: at 60: 1 yr sick; then die	**3.13 (1.46, 5.01)**	3.12 (1.45, 4.95)	3.1 (1.39, 4.91)	3.24 (1.49, 5.1)
Now 35: at 70: 1 yr sick; then die	**2.54 (.87, 4.28)**	2.53 (.87, 4.26)	2.52 (0.87, 4.15)	2.55 (.9, 4.28)
Now 35: at 80: 1 yr sick; then die	**1.42 (.34, 2.53)**	1.41 (.34, 2.51)	1.4 (0.34, 2.44)	1.39 (.3, 2.52)
Now 35: now: 5 yrs sick; then die	**9.66 (5.12, 14.86)**	9.58 (5.04, 14.82)	9.39 (4.83, 14.71)	9.71 (5.01, 14.98)
Now 35: at 40: 5 yrs sick; then die	**8.49 (4.88, 12.61)**	8.4 (4.83, 12.53)	8.21 (4.68, 12.19)	8.61 (5.1, 12.55)

Now 35: at 50: 5 yrs sick; then die	**6.48**	6.41	6.25	6.61
	(4.46, 8.88)	(4.4, 8.77)	(4.26, 8.45)	(4.55, 8.89)
Now 35: at 60: 5 yrs sick; then die	**4.77**	4.71	4.59	4.84
	(3.22, 6.6)	(3.15, 6.52)	(3.02, 6.31)	(3.26, 6.68)
Now 35: at 70: 5 yrs sick; then die	**3.1**	3.06	2.97	3.09
	(1.67, 4.67)	(1.65, 4.61)	(1.53, 4.39)	(1.63, 4.67)
Now 35: at 80: 5 yrs sick; then die	**1.04**	1.01	0.95	1.02
	(.47, 1.63)	(.45, 1.6)	(0.39, 1.5)	(.46, 1.62)
Now 65: Sudden death now	**5.91**	5.88	5.83	5.8
	(1.61, 10.24)	(1.59, 10.18)	(1.57, 10.1)	(1.37, 10.06)
Now 65: Sudden death at 70	**4.37**	4.35	4.32	4.4
	(1.4, 7.29)	(1.41, 7.29)	(1.3, 7.42)	(1.26, 7.48)
Now 65: Sudden death at 80	**1.73**	1.73	1.75	1.81
	(-.07, 3.49)	(-.05, 3.47)	(-0.06, 3.71)	(-.06, 3.88)
Now 65: Sudden death at 90	**-.05**	-.04	-0.03	-.06
	(-1.16, 1.02)	(-1.16, 1.01)	(-1.1, 1.04)	(-1.18, 1.03)
Now 65: now: 1 yr sick; recov	**3.83**	3.82	3.79	3.84
	(1.7, 6.24)	(1.69, 6.2)	(1.58, 6.15)	(1.61, 6.25)
Now 65: at 70: 1 yr sick; recov	**3.33**	3.32	3.3	3.33
	(1.42, 5.44)	(1.43, 5.41)	(1.37, 5.38)	(1.38, 5.47)
Now 65: at 80: 1 yr sick; recov	**2.14**	2.13	2.11	2.13
	(.86, 3.52)	(.86, 3.48)	(0.8, 3.53)	(.8, 3.56)
Now 65: now: 5 yrs sick; recov	**3.85**	3.83	3.77	3.87
	(1.77, 6.07)	(1.75, 6.03)	(1.66, 6.07)	(1.68, 6.21)
Now 65: at 70: 5 yrs sick; recov	**3.28**	3.24	3.16	3.29
	(1.51, 5.16)	(1.48, 5.12)	(1.38, 5.12)	(1.45, 5.3)
Now 65: at 80: 5 yrs sick; recov	**1.81**	1.77	1.66	1.82
	(.77, 2.9)	(.73, 2.88)	(0.66, 2.86)	(.77, 3.04)
Now 65: now: 1 yr sick; then die	**3.67**	3.62	3.51	3.58
	(-.45, 8.06)	(-.49, 7.98)	(-0.64, 7.67)	(-.62, 7.87)
Now 65: at 70: 1 yr sick; then die	**2.73**	2.69	2.6	2.76
	(-.04, 5.45)	(-.06, 5.41)	(-0.17, 5.41)	(-.13, 5.64)
Now 65: at 80: 1 yr sick; then die	**.98**	.96	0.92	1.05
	(-.51, 2.54)	(-.53, 2.5)	(-0.56, 2.53)	(-.46, 2.77)
Now 65: now: 5 yrs sick; then die	**.15**	.12	-0.03	.14
	(-4.45, 4.52)	(-4.46, 4.47)	(-0.68, 0.6)	(-4.5, 5.01)
Now 65: at 70: 5 yrs sick; then die	**-.05**	-.09	0.04	.01
	(-2.98, 2.9)	(-3.02, 2.84)	(-4.55, 4.61)	(-2.9, 3.04)
Now 65: at 80: 5 yrs sick; then die	**-.28**	-.32	-0.19	-.25
	(-1.78, 1.16)	(-1.83, 1.11)	(-3.04, 2.77)	(-1.72, 1.23)
Age profiles				
Now 25: at 25: sudden death	**-10.54**	-10.48	-10.37	-10.7
	(-22.1, .32)	(-21.92, .35)	(-21.75, 0.43)	(-22.34, .21)
Now 30: at 30: sudden death	**-4.23**	-4.19	-4.11	-4.32

	(-11.74, 3.3)	(-11.69, 3.28)	(-11.66, 3.24)	(-12.03, 3.08)
Now 35: at 35: sudden death	**.72**	.73	0.78	.66
	(-4.3, 5.92)	(-4.27, 5.95)	(-4.09, 5.91)	(-4.52, 5.76)
Now 40: at 40: sudden death	**4.37**	4.37	4.39	4.34
	(.42, 8.38)	(.44, 8.37)	(0.42, 8.5)	(.4, 8.39)
Now 45: at 45: sudden death	**6.74**	6.73	6.73	6.73
	(3.12, 10.68)	(3.13, 10.65)	(3.14, 10.49)	(3.05, 10.76)
Now 50: at 50: sudden death	**7.98**	7.96	7.94	7.96
	(4.37, 11.92)	(4.35, 11.99)	(4.2, 11.72)	(4.12, 12.13)
Now 55: at 55: sudden death	**8.15**	8.12	8.1	8.11
	(4.44, 12.2)	(4.38, 12.16)	(4.3, 11.73)	(4.26, 12.17)
Now 60: at 60: sudden death	**7.39**	7.36	7.32	7.32
	(3.59, 11.26)	(3.56, 11.21)	(3.43, 11.08)	(3.38, 11.39)
Now 65: at 65: sudden death	**5.91**	5.88	5.83	5.8
	(1.61, 10.24)	(1.59, 10.18)	(1.57, 10.1)	(1.37, 10.06)
Now 70: at 70: sudden death	**4**	3.97	3.91	3.85
	(-1.59, 9.24)	(-1.61, 9.23)	(-1.64, 9.2)	(-1.86, 9.37)
Now 75: at 75: sudden death	**2.05**	2.02	1.96	1.86
	(-5.46, 8.98)	(-5.48, 8.9)	(-5.53, 8.76)	(-5.86, 8.99)
Now 80: at 80: sudden death	**.16**	.13	0.07	-.09
	(-9.58, 9.16)	(-9.56, 9.15)	(-9.62, 8.73)	(-10.04, 8.89)
Now 25: at 35: 5 yrs sick then die	**8.31**	8.22	8.01	8.42
	(3.36, 13.66)	(3.26, 13.64)	(3.23, 13.49)	(3.41, 13.58)
Now 30: at 40: 5 yrs sick then die	**8.06**	7.98	7.79	8.19
	(4.54, 12.11)	(4.48, 12.01)	(4.36, 11.71)	(4.7, 12.12)
Now 35: at 45: 5 yrs sick then die	**7.43**	7.35	7.18	7.58
	(4.75, 10.48)	(4.68, 10.39)	(4.53, 10.1)	(4.91, 10.5)
Now 40: at 50: 5 yrs sick then die	**6.44**	6.37	6.21	6.59
	(4.25, 8.98)	(4.18, 8.95)	(4.06, 8.5)	(4.49, 9.01)
Now 45: at 55: 5 yrs sick then die	**5.12**	5.05	4.91	5.26
	(3.21, 7.34)	(3.15, 7.23)	(2.94, 6.97)	(3.2, 7.42)
Now 50: at 60: 5 yrs sick then die	**3.63**	3.58	3.44	3.76
	(1.84, 5.66)	(1.8, 5.54)	(1.59, 5.36)	(1.88, 5.75)
Now 55: at 65: 5 yrs sick then die	**2.08**	2.04	1.91	2.2
	(.31, 3.9)	(.28, 3.83)	(0.13, 3.73)	(.4, 4.02)
Now 60: at 70: 5 yrs sick then die	**.69**	.65	0.54	.78
	(-1.03, 2.44)	(-1.05, 2.39)	(-1.13, 2.25)	(-.97, 2.54)
Now 65: at 75: 5 yrs sick then die	**-.22**	-.26	-0.36	-.15
	(-2.03, 1.58)	(-2.08, 1.54)	(-2.13, 1.47)	(-2.03, 1.7)
Now 70: at 80: 5 yrs sick then die	**-.18**	-.21	-0.31	-.13
	(-2.16, 1.69)	(-2.19, 1.65)	(-2.34, 1.64)	(-2.16, 1.91)
Now 75: at 85: 5 yrs sick then die	**1.04**	1	0.9	1.08
	(-1.38, 3.41)	(-1.45, 3.37)	(-1.55, 3.41)	(-1.36, 3.65)
Now 80: at 90: 5 yrs sick then die	**2.9**	2.86	2.74	2.93
	(-.23, 5.96)	(-.27, 5.94)	(-0.39, 6.06)	(-.18, 6.26)
Now 25: ill 6 mo die 6 mo early	**.56**	.56	0.55	.55
	(.3, .85)	(.3, .84)	(0.3, 0.82)	(.3, .81)
Now 30: ill 6 mo die 6 mo early	**.44**	.43	0.43	.43

	(.2, .68)	(.2, .68)	(0.19, 0.67)	(.19, .67)
Now 35: ill 6 mo die 6 mo early	**.26**	.26	0.26	.25
	(.05, .48)	(.05, .48)	(0.05, 0.46)	(.04, .47)
Now 40: ill 6 mo die 6 mo early	**.08**	.08	0.07	.07
	(-.14, .31)	(-.14, .3)	(-0.14, 0.29)	(-.15, .31)
Now 45: ill 6 mo die 6 mo early	**-.11**	-.11	-0.12	-.12
	(-.4, .16)	(-.4, .16)	(-0.39, 0.15)	(-.4, .15)
Now 50: ill 6 mo die 6 mo early	**-.27**	-.27	-0.28	-.28
	(-.64, .06)	(-.64, .06)	(-0.64, 0.05)	(-.64, .05)
Now 55: ill 6 mo die 6 mo early	**-.35**	-.35	-0.36	-.36
	(-.81, .05)	(-.81, .05)	(-0.8, 0.04)	(-.8, .04)
Now 60: ill 6 mo die 6 mo early	**-.29**	-.29	-0.3	-.3
	(-.84, .2)	(-.83, .2)	(-0.82, 0.21)	(-.83, .22)
Now 65: ill 6 mo die 6 mo early	**-.02**	-.02	-0.04	-.03
	(-.69, .63)	(-.7, .62)	(-0.69, 0.61)	(-.7, .63)
Now 70: ill 6 mo die 6 mo early	**.55**	.55	0.53	.55
	(-.43, 1.51)	(-.43, 1.51)	(-0.44, 1.51)	(-.44, 1.48)
Now 75: ill 6 mo die 6 mo early	**1.48**	1.46	1.44	1.48
	(.07, 2.96)	(.07, 2.95)	(0.06, 2.92)	(.09, 2.98)
Now 80: ill 6 mo die 6 mo early	**2.74**	2.72	2.68	2.75
	(.68, 4.92)	(.68, 4.87)	(0.66, 4.93)	(.69, 5.01)
Now 25: at 25: 1 year sick; recover	**1.28**	1.26	1.25	1.31
	(-1.45, 4.14)	(-1.45, 4.11)	(-1.45, 3.98)	(-1.37, 3.98)
Now 30: at 30: 1 year sick; recover	**1.63**	1.62	1.6	1.66
	(-.82, 4.2)	(-.83, 4.19)	(-0.84, 4.07)	(-.8, 4.18)
Now 35: at 35: 1 year sick; recover	**1.92**	1.91	1.89	1.94
	(-.29, 4.27)	(-.3, 4.26)	(-0.35, 4.21)	(-.29, 4.32)
Now 40: at 40: 1 year sick; recover	**2.19**	2.18	2.15	2.21
	(.22, 4.37)	(.21, 4.31)	(0.19, 4.33)	(.2, 4.44)
Now 45: at 45: 1 year sick; recover	**2.42**	2.41	2.38	2.43
	(.51, 4.49)	(.51, 4.47)	(0.52, 4.44)	(.53, 4.53)
Now 50: at 50: 1 year sick; recover	**2.66**	2.64	2.62	2.67
	(.76, 4.67)	(.76, 4.66)	(0.75, 4.65)	(.74, 4.77)
Now 55: at 55: 1 year sick; recover	**2.94**	2.93	2.9	2.95
	(.96, 5.01)	(.94, 5)	(0.93, 5.02)	(.94, 5.11)
Now 60: at 60: 1 year sick; recover	**3.31**	3.3	3.27	3.32
	(1.27, 5.52)	(1.26, 5.52)	(1.21, 5.5)	(1.26, 5.55)
Now 65: at 65: 1 year sick; recover	**3.83**	3.82	3.79	3.84
	(1.7, 6.24)	(1.69, 6.2)	(1.58, 6.15)	(1.61, 6.25)
Now 70: at 70: 1 year sick; recover	**4.57**	4.55	4.52	4.58
	(2.22, 7.19)	(2.21, 7.16)	(2.19, 7.07)	(2.13, 7.21)
Now 75: at 75: 1 year sick; recover	**5.62**	5.6	5.56	5.63
	(3.05, 8.68)	(3.05, 8.62)	(3, 8.41)	(2.92, 8.59)
Now 80: at 80: 1 year sick; recover	**7.00**	6.97	6.93	7.02
	(3.99, 10.56)	(3.97, 10.54)	(3.9, 10.32)	(3.87, 10.53)
Income effects				
Now 45: Sudden death at 45 for income	**$ 2.70**	2.69	2.69	3.05
$10K	**(1.07, 4.46)**	(1.07, 4.45)	(1.07, 4.38)	(1.4, 4.85)

Now 45: Sudden death at 45; Y= $20K	**4.17 (1.77, 6.77)**	4.16 (1.78, 6.75)	4.16 (1.78, 6.65)	4.48 (2.04, 7.14)
Now 45: Sudden death at 45; Y= $30K	**5.41 (2.41, 8.68)**	5.4 (2.41, 8.65)	5.4 (2.42, 8.52)	5.59 (2.54, 8.94)
Now 45: Sudden death at 45; Y= $40K	**6.53 (3.01, 10.36)**	6.52 (3.02, 10.33)	6.52 (3.02, 10.18)	6.55 (2.97, 10.47)
Now 45: Sudden death at 45; Y= $50K	**7.57 (3.58, 11.91)**	7.56 (3.59, 11.87)	7.56 (3.6, 11.7)	7.4 (3.35, 11.84)
Now 45: Sudden death at 45; Y= $60K	**8.55 (4.15, 13.36)**	8.54 (4.16, 13.32)	8.54 (4.16, 13.13)	8.17 (3.69, 13.09)
Now 45: Sudden death at 45; Y= $70K	**9.49 (4.69, 14.73)**	9.48 (4.71, 14.68)	9.48 (4.71, 14.48)	8.89 (4.01, 14.25)
Now 45: Sudden death at 45; Y= $80K	**10.4 (5.23, 16.04)**	10.38 (5.24, 15.99)	10.39 (5.25, 15.76)	9.56 (4.31, 15.34)
Now 45: Sudden death at 45; Y= $90K	**11.28 (5.76, 17.3)**	11.26 (5.78, 17.24)	11.26 (5.78, 17)	10.2 (4.58, 16.36)
Now 45: Sudden death at 45; Y= $100K	**12.13 (6.28, 18.51)**	12.11 (6.3, 18.46)	12.12 (6.31, 18.2)	10.8 (4.85, 17.33)
Now 45: Sudden death at 45; Y= $110K	**12.96 (6.8, 19.69)**	12.95 (6.82, 19.63)	12.95 (6.83, 19.36)	11.37 (5.1, 18.26)
Now 45: Sudden death at 45; Y= $120K	**13.78 (7.31, 20.84)**	13.76 (7.33, 20.78)	13.76 (7.34, 20.5)	11.92 (5.34, 19.15)
Now 45: Sudden death at 45; Y= $130K	**14.58 (7.82, 21.96)**	14.56 (7.84, 21.89)	14.56 (7.85, 21.6)	12.45 (5.57, 20.01)
Now 45: Sudden death at 45; Y= $140K	**15.36 (8.32, 23.05)**	15.34 (8.34, 22.98)	15.35 (8.35, 22.68)	12.96 (5.79, 20.84)
Now 45: Sudden death at 45; Y= $150K	**16.14 (8.82, 24.12)**	16.11 (8.84, 24.05)	16.12 (8.85, 23.74)	13.45 (6.01, 21.64)
Now 45: Sudden death at 45; Y= $160K	**16.9 (9.32, 25.18)**	16.87 (9.34, 25.1)	16.88 (9.35, 24.77)	13.93 (6.21, 22.41)

[b] These estimates are based on an arbitrarily specified initial risk of 0.040 and a risk reduction to 0.036. The risk reduction is thus of magnitude 0.004 and this average WTP for a microrisk reduction is calculated across these 4000 microrisks. Results are minimally different for other baseline and reduced risks. See Table 5-17. The slight curvature of the utility function means that the average WTP amounts will differ somewhat with the size of the risk reduction over which they are calculated. Also, negative point estimates of WTP are not precluded by the formulas used to calculate them. Extreme draws from the estimated joint distribution of the utility parameters can produce negative simulated values. Since there was no opportunity for anyone to express a negative willingness to pay for any of the risk-reduction programs, we adopt a Tobit-like interpretation of the fitted WTP values and interpret negative fitted values as zero.

Table 5-17 Average WTP per μr reduction for different conditions

(See Section 5.6.12. All models *estimated* with $\gamma_3 = 0$, i.e. assuming private risk reduction. Public risk reductions involve payment of program costs even if the individual gets sick, so that $\gamma_3 = 1$. Utility parameters are based on Model 2 in Table 1 of the flagship paper, based on 11,385 choices, with a maximized log-likelihood of -14,841.337. WTP per microrisk reduction depends on baseline risk and size of risk reduction when the Model has curvature with respect to net income, as in our Box-Cox specification.)

Illness profile: age 45 now; …at 45:		Private risk reduction: ($1\ \mu r$; base=0.004)	**Public** risk reduction: ($1\ \mu r$; base=0.004)	**Public** risk reduction: (**1000** μr; base =0.004)	**Public** risk reduction: (**3000** μr; base=0.004)
1.	Sudden death	$ 6.79 (3.2, 10.82)	$ 6.79 (3.2, 10.82)	$ 6.78 (3.13, 10.77)	$ 6.74 (3.12, 10.68)
2.	1 yr sick; nonfatal	2.50 (.66, 4.59)	2.43 (.47, 4.59)	2.43 (.51, 4.5)	2.42 (.51, 4.49)
3.	5 yrs sick; nonfatal	3.08 (1.25, 5.17)	3.07 (1.25, 5.18)	3.06 (1.15, 5.09)	3.05 (1.15, 5.07)
4.	1 yr sick; then die	8.08 (4.57, 11.8)	8.18 (4.57, 12)	8.14 (4.61, 11.93)	8.09 (4.59, 11.82)
5.	5 yrs sick; then die	9.24 (5.35, 13.74)	9.23 (5.36, 13.76)	9.15 (5.35, 13.57)	9.09 (5.33, 13.44)

Illness profile: age 45 now; …at 45:		**Public** risk reduction: ($1\ \mu r$; **base=0.04**)	**Public** risk reduction: (**2000** μr; **base=0.04**)	**Public** risk reduction: (**4000** μr; **base =0.04**)	Private risk reduction: (**4000** μr; **base=0.04**)
1.	Sudden death	$ 6.72 (3.1, 10.66)	$ 6.76 (3.13, 10.72)	$ 6.72 (3.12, 10.64)	$ 6.72 (3.12, 10.64)
2.	1 yr sick; nonfatal	2.37 (0.47, 4.40)	2.42 (0.51, 4.49)	2.42 (0.51, 4.48)	2.42 (0.51, 4.48)
3.	5 yrs sick; nonfatal	3.17 (1.25, 5.19)	3.06 (1.15, 5.08)	3.05 (1.14, 5.06)	3.05 (1.14, 5.06)
4.	1 yr sick; then die	8.22 (4.62, 12.01)	8.11 (4.6, 11.87)	8.06 (4.59, 11.77)	8.06 (4.59, 11.77)
5.	5 yrs sick; then die	9.28 (5.42, 13.63)	9.12 (5.34, 13.5)	9.05 (5.32, 13.37)	9.05 (5.32, 13.37)

Notes: Units are in 2003 US dollars per microrisk reduction for each of five arbitrarily selected illness profiles (rows). Entries reflect 1000 random draws from the joint distribution of estimated parameters. We report the mean, 5th and 95th percentiles for the sampling distribution of calculated *WTP*. Income is set at $42,000.

Table 5-18 Parameters and WTP estimates for different "baseline" choice numbers

Table shows how preferences apparently change systematically from choice to choice across the five choice scenarios (Numbers of choices of each type are 1=2365; 2=2348; 3=2344; 4=2329; 5=2331).

	Model 0 no timedev, no basechoices	Model 1 Base choice =1	Model 2 Base choice=2	Model 3 Base choice=3	Model 4 Base choice=4	Model 5 Base choice=5
$\left(Y_i - c_i^j\right)^{(0.45)} cterm_i^j - \left(Y_i\right)^{(0.45)} yterm_i^j$	.01278	.01317	.01263	.01136	.01199	.01232
	(8.94)***	(7.90)***	(7.89)***	(7.03)***	(7.45)***	(7.35)***
$\Delta\Pi_i^{jS} \log\left(pdvi_i^j + 1\right)$	-41.92	-22.27	-31.54	-42.67	-47.9	-54.68
	(4.78)***	(1.64)	(3.12)***	(4.84)***	(4.59)***	(3.88)***
$\Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	48.72	46.39	46.07	44.76	45.42	45.95
	(1.40)	(1.22)	(1.21)	(1.18)	(1.20)	(1.21)
… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvr_i^j + 1\right)$	-1.619	-1.439	-1.434	-1.400	-1.407	-1.419
	(2.49)**	(1.96)**	(1.95)*	(1.91)*	(1.92)*	(1.93)*
$\Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-1704	-1923	-1929	-1922	-1905	-1906
	(4.09)***	(4.57)***	(4.58)***	(4.57)***	(4.53)***	(4.53)***
… $age_{i0} \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	66.63	73.6	73.82	73.5	72.78	72.85
	(3.93)***	(4.29)***	(4.31)***	(4.29)***	(4.25)***	(4.25)***
… $age_{i0}^2 \times \Delta\Pi_i^{jS} \log\left(pdvl_i^j + 1\right)$	-.6157	-.6429	-.6546	-.6621	-.661	-.669
	(3.70)***	(3.83)***	(3.90)***	(3.94)***	(3.93)***	(3.97)***
$\Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	799.2	873	876	870.6	860.8	862.2
	(3.73)***	(4.05)***	(4.07)***	(4.04)***	(4.00)***	(4.00)***
… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	-31.7	-34.03	-34.16	-33.91	-33.48	-33.54
	(3.60)***	(3.85)***	(3.86)***	(3.83)***	(3.78)***	(3.79)***
… $age_{i0}^2 \times \Delta\Pi_i^{jS} \left[\log\left(pdvl_i^j + 1\right)\right]^2$	.2901	.2973	.303	.305	.3038	.3081
	(3.33)***	(3.40)***	(3.46)***	(3.48)***	(3.46)***	(3.51)***
$\Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right]$	-173.9	-135	-160.8	-185.4	-215.3	-243.4
	(2.96)***	(2.12)**	(2.65)***	(3.10)***	(3.52)***	(3.76)***
$\times \left[\log\left(pdvl_i^j + 1\right)\right]$						

… $age_{i0} \times \Delta\Pi_i^{jS} \left[\log\left(pdvi_i^j + 1\right)\right] \times \left[\log\left(pdvl_i^j + 1\right)\right]$	3.636 (3.26)***	2.744 (2.23)**	3.323 (2.85)***	3.904 (3.40)***	4.505 (3.85)***	5.093 (4.12)***
Scenario adjustment variables (raw acronyms)						
noprogram	.009072	.1229	.08269	.04136	.002479	-.03671
	(0.21)	(2.33)**	(1.77)*	(0.91)	(0.05)	(0.66)
b7term_bn	-.009792	-.009479	-.009354	-.009128	-.009435	-.009546
	(1.93)*	(1.86)*	(1.84)*	(1.80)*	(1.85)*	(1.87)*
dilog_bn	394.6	388	387.4	386.3	387.6	388.1
	(7.70)***	(7.61)***	(7.61)***	(7.58)***	(7.60)***	(7.61)***
dllog_agenow_bn	9.599	9.621	9.624	9.627	9.619	9.618
	(4.20)***	(4.21)***	(4.21)***	(4.21)***	(4.21)***	(4.21)***
dllog2_bn	967.1	925.8	925.8	925.7	925.6	925.7
	(10.49)***	(10.04)***	(10.04)***	(10.06)***	(10.06)***	(10.05)***
dllog2_agenow_bn	-14.94	-14.4	-14.4	-14.4	-14.4	-14.4
	(8.04)***	(7.73)***	(7.73)***	(7.74)***	(7.76)***	(7.76)***
didllog_bn	-667.9	-639.5	-638.1	-635.6	-639	-640.2
	(4.44)***	(4.28)***	(4.27)***	(4.26)***	(4.27)***	(4.28)***
didllog_agenow_bn	6.458	5.992	5.971	5.924	5.973	5.993
	(2.53)**	(2.37)**	(2.36)**	(2.34)**	(2.36)**	(2.37)**
dllog_logldpos	712.4	771.1	775.1	771.7	759.5	759.9
	(2.78)***	(2.99)***	(3.00)***	(2.99)***	(2.94)***	(2.94)***
dllog_agenow_logldpos	-30.18	-33.33	-33.5	-33.36	-32.83	-32.85
	(2.74)***	(2.99)***	(3.01)***	(3.00)***	(2.95)***	(2.95)***
dllog_agenow2_logldpos	.3017	.3372	.3389	.3374	.3321	.3323
	(2.63)***	(2.91)***	(2.93)***	(2.92)***	(2.87)***	(2.87)***
dllog2_logldpos	-339.8	-366.1	-368	-365.7	-359.8	-360.3
	(2.55)**	(2.73)***	(2.75)***	(2.73)***	(2.68)***	(2.69)***
dllog2_agenow_logldpos	14.02	15.34	15.43	15.33	15.06	15.07
	(2.44)**	(2.65)***	(2.66)***	(2.64)***	(2.60)***	(2.60)***

dllog2_agenow2_logldpos	-.1367	-.1512	-.1521	-.1512	-.1483	-.1485
	(2.28)**	(2.50)**	(2.52)**	(2.50)**	(2.45)**	(2.45)**
didllog_logldpos	105.6	110.2	109.7	109	110.3	110.6
	(2.77)***	(2.88)***	(2.87)***	(2.85)***	(2.88)***	(2.89)***
didllog_agenow_logldpos	-2.061	-2.176	-2.167	-2.153	-2.176	-2.183
	(2.66)***	(2.79)***	(2.78)***	(2.76)***	(2.79)***	(2.80)***
drlog_logldneg	11.61	13.34	13.36	13.27	13.2	13.23
	(1.54)	(1.73)*	(1.74)*	(1.72)*	(1.72)*	(1.72)*
dllog_logldneg	592.4	594.6	595.5	590.5	586.8	588.4
	(3.14)***	(3.13)***	(3.14)***	(3.11)***	(3.09)***	(3.10)***
dllog_agenow_logldneg	-25.95	-26.21	-26.25	-26.01	-25.86	-25.94
	(3.36)***	(3.37)***	(3.37)***	(3.34)***	(3.32)***	(3.33)***
dllog_agenow2_logldneg	.2576	.2614	.2618	.2591	.2575	.2584
	(3.40)***	(3.42)***	(3.43)***	(3.39)***	(3.37)***	(3.38)***
dllog2_logldneg	-303.7	-307.9	-308.4	-305.1	-302.9	-303.9
	(3.15)***	(3.17)***	(3.18)***	(3.14)***	(3.12)***	(3.13)***
dllog2_agenow_logldneg	12.98	13.21	13.24	13.08	12.98	13.03
	(3.25)***	(3.29)***	(3.30)***	(3.26)***	(3.23)***	(3.24)***
dllog2_agenow2_logldneg	-.1278	-.1307	-.1309	-.1291	-.1281	-.1287
	(3.22)***	(3.28)***	(3.28)***	(3.24)***	(3.21)***	(3.23)***
didllog_logldneg	56.03	59.64	59.43	58.97	59.49	59.7
	(2.12)**	(2.25)**	(2.24)**	(2.22)**	(2.24)**	(2.25)**
didllog_agenow_logldneg	-.9699	-1.029	-1.026	-1.015	-1.022	-1.027
	(1.90)*	(2.01)**	(2.00)**	(1.98)**	(1.99)**	(2.00)**
dllog_shortwrong	251.4	240.7	240.2	239.8	241	241.3
	(2.95)***	(2.80)***	(2.79)***	(2.78)***	(2.80)***	(2.80)***
dllog_agenow_shortwrong	-5.42	-5.112	-5.083	-5.005	-5.07	-5.103
	(3.27)***	(3.05)***	(3.03)***	(2.98)***	(3.02)***	(3.04)***
dllog2_shortwrong	-133.7	-123.1	-123.1	-123.2	-123.2	-123.2
	(2.91)***	(2.65)***	(2.65)***	(2.66)***	(2.66)***	(2.66)***

dllog2_agenow_shortwrong	3.205	2.944	2.938	2.916	2.928	2.937
	(3.53)***	(3.21)***	(3.20)***	(3.18)***	(3.19)***	(3.20)***
noprogram_shortwrong	.1034	.07707	.07678	.07595	.07659	.07694
	(2.47)**	(1.79)*	(1.78)*	(1.76)*	(1.78)*	(1.79)*
drlog_nocomprisk	-113.3	-155.7	-154.7	-155.1	-158	-158.1
	(1.41)	(1.89)*	(1.88)*	(1.88)*	(1.91)*	(1.91)*
drlog_agenow_nocomprisk	2.314	2.897	2.875	2.836	2.891	2.908
	(1.46)	(1.78)*	(1.77)*	(1.74)*	(1.78)*	(1.79)*
dllog_nocomprisk	412	432.6	433.9	438.6	435.9	434
	(3.11)***	(3.24)***	(3.25)***	(3.29)***	(3.27)***	(3.25)***
dllog_agenow_nocomprisk	-14.6	-16.26	-16.32	-16.53	-16.41	-16.32
	(2.76)***	(3.05)***	(3.06)***	(3.10)***	(3.08)***	(3.06)***
dllog_agenow2_nocomprisk	.1371	.1549	.1556	.1578	.1566	.1557
	(2.74)***	(3.05)***	(3.07)***	(3.11)***	(3.09)***	(3.07)***
b7term_reject	.6451	.6264	.6265	.6269	.6268	.6266
	(21.61)***	(21.18)***	(21.18)***	(21.19)***	(21.19)***	(21.19)***
b7term_affordmiss	-.00814	-.009389	-.009399	-.009418	-.009394	-.009386
	(5.04)***	(5.75)***	(5.75)***	(5.76)***	(5.75)***	(5.75)***
swrdilog	190.3	210.6	211	211.5	210.5	210.2
	(2.23)**	(2.42)**	(2.43)**	(2.43)**	(2.42)**	(2.42)**
drlog_agenow_badtime	-	-3.039	-3.034	-3.108	-3.135	-3.109
		(1.86)*	(1.85)*	(1.90)*	(1.92)*	(1.90)*
noprogram_badtime	-	-.8417	-.8425	-.8412	-.8386	-.8389
		(3.21)***	(3.21)***	(3.20)***	(3.20)***	(3.20)***
drlog_timedev	-	-2.413	-2.429	-2.44	-2.397	-2.39
		(2.49)**	(2.51)**	(2.52)**	(2.47)**	(2.47)**
drlog_agenow_timedev	-	.03155	.03187	.03172	.03077	.03077
		(1.92)*	(1.94)*	(1.93)*	(1.87)*	(1.87)*
dllog_timedev	-	-4.871	-4.863	-4.837	-4.852	-4.863
		(4.13)***	(4.12)***	(4.09)***	(4.11)***	(4.12)***

dllog_agenow_timedev	-	.03565	.0354	.0345	.03501	.03538
		(2.07)**	(2.05)**	(2.00)**	(2.03)**	(2.05)**
dllog2_timedev	-	1.169	1.17	1.169	1.165	1.165
		(3.40)***	(3.40)***	(3.40)***	(3.38)***	(3.38)***
didllog_timedev	-	4.504	4.488	4.491	4.536	4.539
		(2.67)***	(2.66)***	(2.66)***	(2.69)***	(2.69)***
didllog_agenow_timedev	-	-.1698	-.1692	-.1694	-.171	-.1711
		(2.83)***	(2.82)***	(2.83)***	(2.86)***	(2.86)***
didllog_agenow2_timedev	-	.001445	.00144	.001445	.001459	.001458
		(2.73)***	(2.72)***	(2.74)***	(2.76)***	(2.76)***
noprogram_timedev	-	-.01521	-.01521	-.01519	-.01519	-.0152
		(12.22)***	(12.21)***	(12.20)***	(12.20)***	(12.21)***
dllog_timedev2	-	.04798	.04798	.04775	.04771	.04779
		(4.57)***	(4.57)***	(4.54)***	(4.54)***	(4.54)***
dllog_agenow_timedev2	-	-.0004056	-.0004048	-.000397	-.0003977	-.0004005
		(2.75)***	(2.74)***	(2.69)***	(2.69)***	(2.71)***
dllog2_timedev2	-	-.009943	-.009955	-.009997	-.009968	-.009952
		(3.39)***	(3.40)***	(3.41)***	(3.40)***	(3.39)***
noprogram_timedev2	-	.0001025	.0001025	.0001023	.0001024	.0001024
		(9.47)***	(9.47)***	(9.45)***	(9.46)***	(9.46)***
dilog_basechoiceX	-	-10.26	-11.32	-8.869	-5.377	-6.053
		(1.93)*	(2.15)**	(1.74)*	(1.02)	(1.13)
dllog_agenow2_basechoiceX	-	-.01016	-.01111	-.009465	-.006432	-.006839
		(2.51)**	(2.78)***	(2.46)**	(1.62)	(1.70)*
dllog2_agenow2_basechoiceX	-	.004571	.004891	.004336	.003314	.003452
		(2.38)**	(2.56)**	(2.31)**	(1.74)*	(1.80)*
didllog_basechoiceX	-	-26.15	-25.85	-27.15	-28.29	-27.89
		(2.25)**	(2.22)**	(2.33)**	(2.43)**	(2.40)**
didllog_agenow_basechoiceX	-	.5997	.6100	.6072	.5776	.5767
		(2.62)***	(2.66)***	(2.64)***	(2.52)**	(2.52)**

noprogram_basechoiceX	-	-.03985	-.03983	-.03963	-.03964	-.03972
		(2.67)***	(2.67)***	(2.66)***	(2.66)***	(2.66)***
b7term_basechoiceXsq	-	.0001118	.0004102	.001265	.0006141	.00025
		(0.78)	(1.70)*	(3.32)***	(2.51)**	(1.73)*
Number of alternatives	34,155	34,155	34,155	34,155	34,155	34,155
Number of choices	11,385	11,385	11,385	11,385	11,385	11,385
Log L	-14848.463	-14700.61	-14699.453	-14695.346	-14697.722	-14699.407

Table 5-19 WTP estimates for different "baseline" choice numbers

Table shows how WTP estimates change systematically from choice to choice across the five choice scenarios (Numbers of choices of each type are 1=2365; 2=2348; 3=2344; 4=2329; 5=2331).

	Model 0 no timedev, no basechoices	Model 1 Base choice =1	Model 2 Base choice=2	Model 3 Base choice=3	Model 4 Base choice=4	Model 5 Base choice=5
WTP for a microrisk reduction						
Income $42,000						
Now 45: Sudden death at 45	$ 7.40 (3.59, 11.43)	$ 7.91 (3.66, 12.25)	$ 7.65 (3.45, 12.11)	$ 7.78 (3.16, 12.76)	$ 6.77 (2.62, 11.32)	$ 6.04 (1.79, 10.72)
Now 45: at 45: 1 yr sick; recov	2.58 (.51, 4.68)	1.62 (-.45, 3.66)	1.92 (-.22, 4.11)	2.45 (0, 4.9)	2.42 (.07, 4.74)	2.55 (.29, 4.87)
Now 45: at 45: 5 yrs sick; recov	3.63 (1.7, 5.65)	2.13 (.14, 4.19)	2.73 (.49, 4.86)	3.7 (1.44, 6.24)	3.77 (1.51, 6.03)	4.06 (1.74, 6.39)
Now 45: at 45: 1 yr sick; then die	8.39 (4.52, 12.44)	8.48 (4.51, 12.77)	8.45 (4.4, 12.64)	8.89 (4.57, 13.59)	8.13 (4.32, 12.53)	7.60 (3.56, 12.18)
Now 45: at 45: 5 yrs sick; then die	8.93 (4.84, 13.1)	8.25 (4.04, 12.85)	8.63 (4.45, 12.85)	9.56 (5.01, 14.74)	9.28 (4.88, 14.36)	9.15 (4.7, 14.42)
Income=$25,000						
Now 45: Sudden death at 45	5.3 (2.44, 8.32)	5.68 (2.5, 8.93)	5.49 (2.34, 8.82)	5.59 (2.12, 9.31)	4.83 (1.72, 8.24)	4.28 (1.09, 7.79)
Now 45: at 45: 1 yr sick; recov	1.94 (.38, 3.51)	1.21 (-.33, 2.75)	1.45 (-.16, 3.09)	1.84 (0, 3.68)	1.82 (.06, 3.56)	1.92 (.22, 3.66)
Now 45: at 45: 5 yrs sick; recov	2.73 (1.28, 4.24)	1.60 (.1, 3.15)	2.05 (.37, 3.65)	2.78 (1.08, 4.69)	2.83 (1.13, 4.53)	3.05 (1.31, 4.8)
Now 45: at 45: 1 yr sick; then die	6.06 (3.16, 9.09)	6.12 (3.14, 9.33)	6.1 (3.07, 9.24)	6.43 (3.2, 9.94)	5.86 (3.01, 9.15)	5.46 (2.44, 8.89)
Now 45: at 45: 5 yrs sick; then die	6.51 (3.44, 9.63)	6.00 (2.84, 9.44)	6.29 (3.15, 9.44)	6.98 (3.58, 10.85)	6.77 (3.47, 10.57)	6.67 (3.34, 10.61)
Income= $67,500						
Now 45: Sudden death at 45	10.12 (5.16, 15.37)	10.78 (5.25, 16.43)	10.45 (4.98, 16.25)	10.62 (4.6, 17.1)	9.30 (3.9, 15.22)	8.35 (2.83, 14.44)
Now 45: at 45: 1 yr sick; recov	3.35 (.65, 6.07)	2.09 (-.58, 4.75)	2.49 (-.29, 5.33)	3.18 (-.01, 6.36)	3.14 (.09, 6.15)	3.31 (.37, 6.32)

Now 45: at 45: 5 yrs sick; recov	4.71 (2.2, 7.34)	2.77 (.17, 5.44)	3.54 (.63, 6.32)	4.80 (1.86, 8.11)	4.89 (1.95, 7.84)	5.27 (2.26, 8.31)
Now 45: at 45: 1 yr sick; then die	11.38 (6.35, 16.66)	11.5 (6.32, 17.09)	11.46 (6.19, 16.92)	12.03 (6.41, 18.15)	11.04 (6.09, 16.77)	10.35 (5.1, 16.32)
Now 45: at 45: 5 yrs sick; then die	11.98 (6.66, 17.42)	11.11 (5.61, 17.09)	11.6 (6.15, 17.09)	12.81 (6.89, 19.56)	12.44 (6.71, 19.06)	12.27 (6.48, 19.15)
Now 35: Sudden death now	.89 (-4.92, 6.23)	1.22 (-4.22, 6.55)	.89 (-4.82, 6.14)	.44 (-5.61, 6.56)	.12 (-5.23, 5.94)	-.24 (-5.83, 5.73)
Now 35: Sudden death at 40	1.43 (-3.22, 5.57)	1.74 (-2.66, 6.02)	1.59 (-2.77, 5.73)	1.42 (-3.26, 6.23)	1.12 (-3.16, 5.74)	.84 (-3.56, 5.49)
Now 35: Sudden death at 50	2.29 (-.49, 4.89)	2.56 (-.07, 5.21)	2.67 (0, 5.35)	2.93 (.04, 6.04)	2.68 (.01, 5.57)	2.54 (-.09, 5.47)
Now 35: Sudden death at 60	2.73 (.82, 4.82)	2.93 (1.01, 4.99)	3.21 (1.15, 5.32)	3.72 (1.4, 6.31)	3.53 (1.39, 5.82)	3.48 (1.31, 5.83)
Now 35: Sudden death at 70	2.58 (.83, 4.52)	2.73 (1.02, 4.56)	3.05 (1.22, 4.89)	3.62 (1.42, 5.84)	3.49 (1.46, 5.74)	3.5 (1.6, 5.79)
Now 35: Sudden death at 80	1.64 (.45, 2.92)	1.71 (.55, 2.92)	1.94 (.67, 3.2)	2.33 (.9, 3.79)	2.27 (.88, 3.79)	2.3 (.99, 3.84)
Now 35: now: 1 yr sick; recov	1.78 (-.44, 4.23)	1.1 (-1.28, 3.41)	1.36 (-1.14, 3.74)	1.81 (-.88, 4.52)	1.79 (-.9, 4.41)	1.91 (-.73, 4.44)
Now 35: at 40: 1 yr sick; recov	1.59 (-.45, 3.81)	.99 (-1.18, 3.09)	1.22 (-1.02, 3.4)	1.61 (-.84, 4.05)	1.59 (-.86, 4)	1.69 (-.7, 3.96)
Now 35: at 50: 1 yr sick; recov	1.24 (-.43, 3.04)	.81 (-.94, 2.47)	.98 (-.86, 2.76)	1.27 (-.72, 3.26)	1.24 (-.69, 3.15)	1.31 (-.59, 3.18)
Now 35: at 60: 1 yr sick; recov	.95 (-.31, 2.32)	.65 (-.66, 1.92)	.77 (-.59, 2.12)	.99 (-.51, 2.51)	.97 (-.48, 2.42)	1.01 (-.39, 2.42)
Now 35: at 70: 1 yr sick; recov	.7 (-.13, 1.59)	.52 (-.33, 1.36)	.61 (-.3, 1.5)	.77 (-.23, 1.78)	.75 (-.21, 1.71)	.78 (-.13, 1.74)
Now 35: at 80: 1 yr sick; recov	.48 (.12, .87)	.39 (.03, .77)	.46 (.05, .87)	.59 (.12, 1.08)	.58 (.15, 1.03)	.6 (.19, 1.07)
Now 35: now: 5 yrs sick; recov	3 (.86, 5.31)	1.74 (-.53, 4)	2.31 (-.14, 4.73)	3.24 (.65, 6.02)	3.32 (.76, 5.87)	3.6 (1, 6.33)

Now 35: at 40: 5 yrs sick; recov	2.7 (.77, 4.81)	1.59 (-.46, 3.64)	2.09 (-.13, 4.27)	2.92 (.6, 5.43)	2.99 (.67, 5.32)	3.24 (.9, 5.7)
Now 35: at 50: 5 yrs sick; recov	2.13 (.6, 3.8)	1.28 (-.33, 2.88)	1.67 (-.07, 3.33)	2.32 (.48, 4.29)	2.37 (.56, 4.22)	2.56 (.73, 4.48)
Now 35: at 60: 5 yrs sick; recov	1.62 (.52, 2.82)	1.01 (-.14, 2.15)	1.3 (.06, 2.52)	1.79 (.45, 3.22)	1.82 (.55, 3.17)	1.96 (.65, 3.37)
Now 35: at 70: 5 yrs sick; recov	1.17 (.5, 1.89)	.76 (.1, 1.46)	.98 (.23, 1.71)	1.34 (.51, 2.23)	1.36 (.58, 2.18)	1.46 (.65, 2.36)
Now 35: at 80: 5 yrs sick; recov	.78 (.49, 1.09)	.55 (.25, .86)	.7 (.41, 1.04)	.96 (.62, 1.37)	.98 (.63, 1.39)	1.05 (.7, 1.47)
Now 35: now: 1 yr sick; then die	4.41 (-.45, 9.18)	3.85 (-.89, 8.89)	4.14 (-.57, 9.03)	4.63 (-.66, 10.15)	4.74 (-.04, 9.73)	4.79 (-.18, 10.29)
Now 35: at 40: 1 yr sick; then die	4.18 (.39, 7.84)	3.78 (.15, 7.61)	4.11 (.57, 7.89)	4.68 (.62, 8.99)	4.72 (.91, 8.72)	4.76 (.98, 8.95)
Now 35: at 50: 1 yr sick; then die	3.88 (1.58, 6.09)	3.71 (1.62, 5.99)	4.12 (1.9, 6.52)	4.82 (2.47, 7.56)	4.76 (2.53, 7.3)	4.8 (2.59, 7.33)
Now 35: at 60: 1 yr sick; then die	3.56 (1.87, 5.43)	3.52 (1.88, 5.36)	3.96 (2.19, 5.91)	4.71 (2.54, 7.03)	4.62 (2.72, 7.03)	4.68 (2.8, 6.85)
Now 35: at 70: 1 yr sick; then die	2.94 (1.26, 4.82)	2.94 (1.36, 4.66)	3.34 (1.61, 5.09)	4.03 (1.96, 6.11)	3.95 (2.02, 6.14)	4.01 (2.19, 6.18)
Now 35: at 80: 1 yr sick; then die	1.66 (.57, 2.81)	1.66 (.62, 2.74)	1.91 (.78, 3.04)	2.32 (1.05, 3.63)	2.28 (1.06, 3.66)	2.33 (1.17, 3.76)
Now 35: now: 5 yrs sick; then die	9.4 (4.22, 14.9)	7.6 (2.33, 13.22)	8.82 (3.5, 14.5)	10.72 (4.9, 16.96)	11.4 (5.61, 17.43)	12.04 (6.34, 18.6)
Now 35: at 40: 5 yrs sick; then die	8.45 (4.62, 12.56)	6.97 (3.04, 11.27)	8.11 (3.97, 12.46)	9.89 (5.39, 14.8)	10.41 (5.92, 15.1)	10.96 (6.62, 15.89)
Now 35: at 50: 5 yrs sick; then die	6.78 (4.6, 9.25)	5.82 (3.61, 8.41)	6.78 (4.41, 9.34)	8.32 (5.68, 11.53)	8.6 (6.01, 11.7)	9 (6.27, 12.24)
Now 35: at 60: 5 yrs sick; then die	5.22 (3.51, 7.17)	4.64 (3.07, 6.37)	5.42 (3.62, 7.33)	6.69 (4.61, 9.14)	6.81 (4.71, 9.47)	7.09 (5.08, 9.49)
Now 35: at 70: 5 yrs sick; then die	3.54 (2.01, 5.17)	3.2 (1.87, 4.63)	3.76 (2.24, 5.39)	4.67 (2.85, 6.67)	4.71 (3.01, 6.75)	4.9 (3.24, 6.92)
Now 35: at 80: 5 yrs sick; then die	1.29	1.08	1.31	1.7	1.72	1.81

	(.69, 1.93)	(.53, 1.62)	(.72, 1.95)	(1.01, 2.46)	(1.05, 2.48)	(1.17, 2.55)
Now 65: Sudden death now	7.1 (2.69, 11.62)	7.84 (3, 12.75)	7.3 (2.79, 12.36)	7.3 (2.32, 12.87)	6.01 (1.12, 11.38)	4.82 (.09, 10.07)
Now 65: Sudden death at 70	5.23 (2.09, 8.59)	4.93 (1.43, 8.45)	4.78 (1.61, 8.36)	5.07 (1.56, 9.06)	4.29 (.84, 8.17)	3.61 (.18, 7.33)
Now 65: Sudden death at 80	2.01 (.09, 3.99)	.22 (-1.87, 2.36)	.66 (-1.44, 2.79)	1.34 (-.97, 3.7)	1.41 (-.79, 3.68)	1.55 (-.75, 3.78)
Now 65: Sudden death at 90	-.11 (-1.27, 1.01)	-1.28 (-2.69, .01)	-.93 (-2.16, .31)	-.54 (-1.89, .83)	-.26 (-1.61, 1.02)	.05 (-1.38, 1.44)
Now 65: now: 1 yr sick; recov	4.51 (2.13, 7)	2.76 (.38, 5.39)	3.24 (.62, 5.91)	4.01 (1.23, 7.05)	3.98 (1.34, 6.82)	4.17 (1.6, 7.05)
Now 65: at 70: 1 yr sick; recov	3.91 (1.79, 6.1)	2.3 (.22, 4.62)	2.74 (.44, 5.13)	3.44 (.97, 6.09)	3.43 (1.09, 5.91)	3.62 (1.34, 6.1)
Now 65: at 80: 1 yr sick; recov	2.5 (1.06, 3.96)	1.21 (-.21, 2.72)	1.57 (0, 3.14)	2.11 (.37, 3.91)	2.17 (.56, 3.85)	2.35 (.76, 4)
Now 65: now: 5 yrs sick; recov	4.82 (2.68, 7.17)	2.72 (.42, 5.08)	3.37 (.99, 5.81)	4.43 (1.84, 7.29)	4.46 (2, 7.14)	4.75 (2.33, 7.52)
Now 65: at 70: 5 yrs sick; recov	4.11 (2.31, 6.08)	2.2 (.25, 4.19)	2.8 (.77, 4.88)	3.74 (1.48, 6.18)	3.8 (1.74, 6.07)	4.08 (1.99, 6.45)
Now 65: at 80: 5 yrs sick; recov	2.34 (1.25, 3.5)	.86 (-.34, 2.04)	1.34 (.13, 2.58)	2.06 (.73, 3.42)	2.19 (.95, 3.55)	2.46 (1.17, 3.83)
Now 65: now: 1 yr sick; then die	4.31 (.27, 8.49)	5.48 (.86, 10.08)	4.5 (.1, 9.17)	3.77 (-1.3, 8.91)	2.35 (-2.43, 7.17)	.96 (-4.02, 5.87)
Now 65: at 70: 1 yr sick; then die	3.2 (.53, 6.03)	3.14 (.11, 6.2)	2.7 (-.11, 5.8)	2.49 (-.72, 5.9)	1.65 (-1.34, 4.85)	.86 (-2.23, 4.07)
Now 65: at 80: 1 yr sick; then die	1.11 (-.53, 2.68)	-.63 (-2.43, 1.14)	-.28 (-2.01, 1.48)	.22 (-1.69, 2.11)	.29 (-1.51, 2.11)	.43 (-1.5, 2.28)
Now 65: at 90: 1 yr sick; then die	.05 (-.61, .71)	-.75 (-1.58, 0)	-.49 (-1.23, .21)	-.19 (-.95, .61)	-	.22 (-.6, 1.02)

Now 65: now: 5 yrs sick; then die	-.03 (-4.49, 4.17)	1.05 (-4.32, 6.08)	-.23 (-5.12, 4.31)	-1.68 (-7.14, 3.44)	-2.98 (-8.35, 2.09)	-4.31 (-10.01, .94)
Now 65: at 70: 5 yrs sick; then die	-.19 (-3.06, 2.51)	-.44 (-3.77, 2.68)	-1.05 (-4.24, 1.85)	-1.74 (-5.3, 1.42)	-2.43 (-5.84, .76)	-3.11 (-6.92, .24)
Now 65: at 80: 5 yrs sick; then die	-.32 (-1.89, 1.08)	-2.2 (-4.11, -.57)	-1.83 (-3.54, -.27)	-1.43 (-3.42, .33)	-1.21 (-3.06, .48)	-.93 (-2.9, .92)
Now 25: at 25: sudden death	-11.01 (-23.53, -.49)	-10.77 (-22.52, .77)	-11.41 (-24.02, -.65)	-13.04 (-27.98, -1.22)	-12.25 (-25.34, -.5)	-12.14 (-25.38, -.32)
Now 30: at 30: sudden death	-4.35 (-12.54, 2.92)	-4.08 (-11.94, 3.44)	-4.53 (-12.65, 2.94)	-5.49 (-15.08, 3.11)	-5.31 (-13.71, 2.87)	-5.44 (-14.23, 2.32)
Now 35: at 35: sudden death	.89 (-4.92, 6.23)	1.22 (-4.22, 6.55)	.89 (-4.82, 6.14)	.44 (-5.61, 6.56)	.12 (-5.23, 5.94)	-.24 (-5.83, 5.73)
Now 40: at 40: sudden death	4.8 (.4, 9)	5.22 (.68, 9.64)	4.95 (.56, 9.35)	4.86 (.14, 9.99)	4.14 (-.25, 8.6)	3.58 (-.77, 8.49)
Now 45: at 45: sudden death	7.4 (3.59, 11.43)	7.91 (3.66, 12.25)	7.65 (3.45, 12.11)	7.78 (3.16, 12.76)	6.77 (2.62, 11.32)	6.04 (1.79, 10.72)
Now 50: at 50: sudden death	8.83 (5.01, 12.83)	9.43 (5.03, 14.06)	9.14 (4.97, 13.65)	9.37 (4.72, 14.41)	8.15 (3.85, 12.89)	7.26 (3.05, 12.3)
Now 55: at 55: sudden death	9.15 (5.4, 13.24)	9.84 (5.57, 14.52)	9.48 (5.51, 14.08)	9.7 (5.2, 14.71)	8.37 (4.09, 13.27)	7.35 (3.02, 12.16)
Now 60: at 60: sudden death	8.51 (4.79, 12.7)	9.25 (4.99, 13.87)	8.8 (4.81, 13.47)	8.94 (4.35, 14.1)	7.59 (3.25, 12.61)	6.46 (2.06, 11.32)
Now 65: at 65: sudden death	7.1 (2.69, 11.62)	7.84 (3, 12.75)	7.3 (2.79, 12.36)	7.3 (2.32, 12.87)	6.01 (1.12, 11.38)	4.82 (.09, 10.07)
Now 70: at 70: sudden death	5.22 (-.14, 10.91)	5.86 (-.26, 11.77)	5.27 (-.46, 11.52)	5.15 (-1.27, 12.08)	3.98 (-2.23, 10.54)	2.78 (-3.18, 9.04)
Now 75: at 75: sudden death	3.3 (-3.99, 10.83)	3.81 (-4.02, 11.33)	3.18 (-4.72, 11.24)	2.96 (-5.53, 11.35)	1.95 (-6.23, 9.97)	.74 (-7.27, 8.8)
Now 80: at 80: sudden death	1.45 (-7.99, 11.25)	1.87 (-8.27, 11.24)	1.19 (-9.18, 11)	.88 (-10.37, 11.96)	.02 (-10.78, 10.18)	-1.22 (-11.71, 8.88)
Now 25: at 35: 5 yrs sick then die	8.54 (3.21, 14.19)	7.1 (1.58, 12.77)	8.71 (3.53, 14.44)	11.12 (5.02, 17.92)	12.05 (6.2, 18.35)	12.89 (7.13, 19.39)
Now 30: at 40: 5 yrs sick then die	8.26	6.94	8.27	10.31	10.95	11.6

	(4.24, 12.47)	(3.12, 11.24)	(4.3, 12.68)	(5.95, 15.46)	(6.54, 15.7)	(7.17, 16.58)
Now 35: at 45: 5 yrs sick then die	7.59	6.39	7.43	9.1	9.49	9.96
	(4.65, 10.95)	(3.5, 9.67)	(4.31, 10.73)	(5.65, 13.04)	(6.1, 13.29)	(6.55, 13.91)
Now 40: at 50: 5 yrs sick then die	6.56	5.48	6.25	7.53	7.71	8
	(4.22, 9.23)	(3.11, 8.15)	(3.76, 8.83)	(4.67, 10.7)	(4.9, 10.83)	(5.15, 11.27)
Now 45: at 55: 5 yrs sick then die	5.2	4.2	4.71	5.6	5.62	5.77
	(3.16, 7.44)	(2.15, 6.33)	(2.64, 6.89)	(3.25, 8.14)	(3.35, 8.17)	(3.41, 8.5)
Now 50: at 60: 5 yrs sick then die	3.66	2.72	2.99	3.52	3.42	3.43
	(1.79, 5.61)	(.86, 4.69)	(1.13, 4.99)	(1.55, 5.73)	(1.44, 5.62)	(1.33, 5.77)
Now 55: at 65: 5 yrs sick then die	2.07	1.12	1.19	1.41	1.24	1.15
	(.38, 3.76)	(-.67, 2.98)	(-.61, 2.93)	(-.54, 3.37)	(-.67, 3.17)	(-.94, 3.17)
Now 60: at 70: 5 yrs sick then die	.62	-.43	-.47	-.47	-.67	-.81
	(-1.07, 2.27)	(-2.29, 1.29)	(-2.38, 1.3)	(-2.63, 1.36)	(-2.71, 1.13)	(-3.03, 1.24)
Now 65: at 75: 5 yrs sick then die	-.32	-1.61	-1.66	-1.72	-1.88	-2.01
	(-2.11, 1.4)	(-3.68, .3)	(-3.73, .19)	(-4.14, .29)	(-4.12, .1)	(-4.42, .16)
Now 70: at 80: 5 yrs sick then die	-.29	-1.97	-1.86	-1.72	-1.8	-1.83
	(-2.34, 1.61)	(-4.27, .21)	(-4.24, .26)	(-4.4, .76)	(-4.58, .53)	(-4.63, .71)
Now 75: at 85: 5 yrs sick then die	.99	-1.07	-.69	-.17	-.19	-.1
	(-1.64, 3.48)	(-4.01, 1.71)	(-3.64, 2.06)	(-3.3, 2.97)	(-3.3, 2.66)	(-3.34, 2.89)
Now 80: at 90: 5 yrs sick then die	2.97	.73	1.37	2.31	2.29	2.47
	(-.38, 6.34)	(-2.88, 4.3)	(-2.51, 4.85)	(-1.6, 6.22)	(-1.31, 6.03)	(-1.12, 6.25)
Now 25: ill 6 mo die 6 mo early	.62	.67	.72	.82	.78	.77
	(.34, .92)	(.4, .97)	(.42, 1.04)	(.49, 1.22)	(.47, 1.16)	(.47, 1.14)
Now 30: ill 6 mo die 6 mo early	.5	.52	.57	.67	.65	.64
	(.24, .77)	(.29, .78)	(.32, .85)	(.37, 1)	(.37, .96)	(.38, .96)
Now 35: ill 6 mo die 6 mo early	.32	.31	.36	.44	.44	.45
	(.1, .54)	(.1, .51)	(.14, .58)	(.2, .7)	(.2, .71)	(.22, .72)
Now 40: ill 6 mo die 6 mo early	.13	.06	.12	.2	.22	.26
	(-.09, .36)	(-.18, .29)	(-.11, .36)	(-.05, .45)	(-.03, .48)	(.01, .53)
Now 45: ill 6 mo die 6 mo early	-.07	-.22	-.15	-.07	-.02	.05
	(-.35, .2)	(-.54, .09)	(-.46, .13)	(-.39, .26)	(-.34, .29)	(-.28, .36)
Now 50: ill 6 mo die 6 mo early	-.24	-.5	-.41	-.32	-.23	-.14
	(-.6, .1)	(-.92, -.1)	(-.81, -.04)	(-.75, .09)	(-.65, .16)	(-.57, .25)
Now 55: ill 6 mo die 6 mo early	-.34	-.73	-.6	-.48	-.36	-.23

	(-.77, .08)	(-1.27, -.24)	(-1.09, -.15)	(-1.03, .04)	(-.89, .13)	(-.77, .27)
Now 60: ill 6 mo die 6 mo early	-.29 (-.83, .22)	-.85 (-1.51, -.25)	-.67 (-1.26, -.1)	-.5 (-1.16, .15)	-.34 (-.98, .28)	-.17 (-.84, .45)
Now 65: ill 6 mo die 6 mo early	-.02 (-.72, .65)	-.79 (-1.63, 0)	-.55 (-1.29, .19)	-.27 (-1.08, .56)	-.08 (-.89, .7)	.12 (-.73, .96)
Now 70: ill 6 mo die 6 mo early	.55 (-.45, 1.55)	-.46 (-1.62, .63)	-.12 (-1.24, .91)	.3 (-.8, 1.5)	.51 (-.62, 1.61)	.75 (-.38, 1.93)
Now 75: ill 6 mo die 6 mo early	1.48 (.04, 3)	.22 (-1.36, 1.86)	.69 (-.89, 2.27)	1.31 (-.26, 3.09)	1.5 (-.1, 3.19)	1.78 (.17, 3.48)
Now 80: ill 6 mo die 6 mo early	2.77 (.69, 5)	1.29 (-.92, 3.61)	1.88 (-.46, 4.22)	2.75 (.44, 5.34)	2.9 (.62, 5.38)	3.19 (.95, 5.64)
Now 45: Sudden death at 45 for income $10K	2.99 (1.28, 4.8)	3.22 (1.31, 5.16)	3.11 (1.21, 5.1)	3.16 (1.08, 5.39)	2.71 (.84, 4.75)	2.38 (.46, 4.48)
Now 45: Sudden death at 45 for income $20K	4.61 (2.08, 7.27)	4.94 (2.13, 7.8)	4.77 (1.99, 7.71)	4.86 (1.79, 8.14)	4.19 (1.44, 7.19)	3.7 (.89, 6.8)
Now 45: Sudden death at 45 for income $30K	5.96 (2.79, 9.3)	6.37 (2.85, 9.97)	6.16 (2.67, 9.86)	6.27 (2.43, 10.4)	5.43 (1.99, 9.2)	4.82 (1.3, 8.71)
Now 45: Sudden death at 45 for income $40K	7.17 (3.46, 11.09)	7.66 (3.53, 11.89)	7.42 (3.32, 11.75)	7.54 (3.04, 12.39)	6.56 (2.52, 10.98)	5.84 (1.71, 10.4)
Now 45: Sudden death at 45 for income $50K	8.3 (4.09, 12.74)	8.85 (4.17, 13.64)	8.57 (3.94, 13.48)	8.72 (3.62, 14.2)	7.6 (3.03, 12.61)	6.79 (2.12, 11.96)
Now 45: Sudden death at 45 for income $60K	9.36 (4.71, 14.27)	9.97 (4.8, 15.27)	9.66 (4.54, 15.1)	9.82 (4.18, 15.9)	8.59 (3.53, 14.14)	7.69 (2.52, 13.41)
Now 45: Sudden death at 45 for income $70K	10.37 (5.31, 15.72)	11.04 (5.4, 16.81)	10.7 (5.12, 16.62)	10.88 (4.73, 17.49)	9.54 (4.03, 15.58)	8.56 (2.93, 14.78)
Now 45: Sudden death at 45 for income $80K	11.35 (5.89, 17.11)	12.06 (6, 18.28)	11.7 (5.69, 18.08)	11.89 (5.28, 19.02)	10.45 (4.51, 16.95)	9.39 (3.33, 16.09)
Now 45: Sudden death at 45 for income $90K	12.29 (6.46, 18.44)	13.05 (6.58, 19.69)	12.67 (6.25, 19.48)	12.87 (5.81, 20.48)	11.33 (5, 18.27)	10.2 (3.73, 17.36)
Now 45: Sudden death at 45 for income $100K	13.2 (7.03, 19.72)	14.01 (7.15, 21.05)	13.61 (6.81, 20.82)	13.82 (6.33, 21.89)	12.18 (5.47, 19.54)	10.99 (4.14, 18.57)
Now 45: Sudden death at 45 for income $110K	14.09 (7.59, 20.97)	14.95 (7.71, 22.37)	14.52 (7.35, 22.13)	14.74 (6.85, 23.25)	13.02 (5.95, 20.78)	11.76 (4.54, 19.76)
Now 45: Sudden death at 45 for income $120K	14.97	15.86	15.41	15.65	13.84	12.52

	(8.14, 22.18)	(8.27, 23.65)	(7.89, 23.4)	(7.37, 24.57)	(6.42, 21.98)	(4.94, 20.91)
Now 45: Sudden death at 45 for income $130K	15.82 (8.68, 23.36)	16.76 (8.82, 24.89)	16.28 (8.42, 24.63)	16.53 (7.88, 25.86)	14.64 (6.88, 23.15)	13.26 (5.34, 22.03)
Now 45: Sudden death at 45 for income $140K	16.66 (9.22, 24.51)	17.63 (9.36, 26.11)	17.14 (8.95, 25.84)	17.4 (8.38, 27.12)	15.43 (7.35, 24.3)	13.99 (5.74, 23.13)
Now 45: Sudden death at 45 for income $150K	17.48 (9.75, 25.64)	18.49 (9.9, 27.3)	17.98 (9.47, 27.02)	18.25 (8.88, 28.35)	16.2 (7.81, 25.42)	14.71 (6.14, 24.2)
Now 45: Sudden death at 45 for income $160K	18.29 (10.28, 26.75)	19.34 (10.44, 28.47)	18.81 (9.99, 28.18)	19.09 (9.38, 29.56)	16.97 (8.27, 26.52)	15.42 (6.54, 25.26)

5.8 Section 5 Figures

Illness Profile 1: Sudden death in the current period (usual VSL illness profile)

Illness Profile	Lost Life Years
Health Status	

Illness Profile 2: A nonfatal illness (with recovery) that reduces life expectancy

Illness Profile	Latency Period	Sick Years	Recovered Years	Lost Life Years
Health Status	healthy	sick	recovered	

Illness Profile 3: A fatal illness (no recovery)

Illness Profile	Latency Period	Sick Years	Lost Life Years
Health Status	healthy	sick	

Figure 5-1 Depiction of alternative illness profiles

```
hausman fixed nofixed;

Note: the rank of the differenced variance matrix (9) does not equal the number of
      coefficients being tested (13); be sure this is what you expect, or there may be
      problems computing the test. Examine the output of your estimators for anything
      unexpected and possibly consider scaling your variables so that the coefficients
      are on a similar scale.

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed       nofixed       Difference          S.E.
-------------+------------------------------------------------------------------
      b7term |    .0138705     .0136397        .0002308        .0001874
       dilog |   -49.15848    -48.05254       -1.105939        .8740273
       drlog |   -16.75242    -17.09082        .3383987               .
       dllog |   -561.9437    -500.5019       -61.44189        43.40268
dllog_agenow |    19.63725     18.28144        1.355803        1.790943
dllog_age~w2 |   -.1800343    -.1764901       -.0035442        .0174983
      dllog2 |    194.5601     175.4035        19.15658        11.55654
dllog2_age~w |   -7.504353    -7.121171       -.3831828        .5158851
dllog2_ag~w2 |    .0714078     .0710271        .0003806        .0053307
     didllog |     104.025     99.29108         4.73395         3.41158
didllog_ag~w |   -4.500657    -4.335806       -.1648511        .1710725
didllog_a~w2 |    .0561213     .0545688        .0015525        .0020757
      sdilog |    3.372028     3.006647        .3653809        .4531743
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from clogit
           B = inconsistent under Ha, efficient under Ho; obtained from clogit

    Test:  Ho:  difference in coefficients not systematic

                  chi2(9) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       13.77
                Prob>chi2 =      0.1307
                (V_b-V_B is not positive definite)
```

Figure 5-2 Hausman test using Stata
Preliminary specification with selection correction but without scenario adjustment

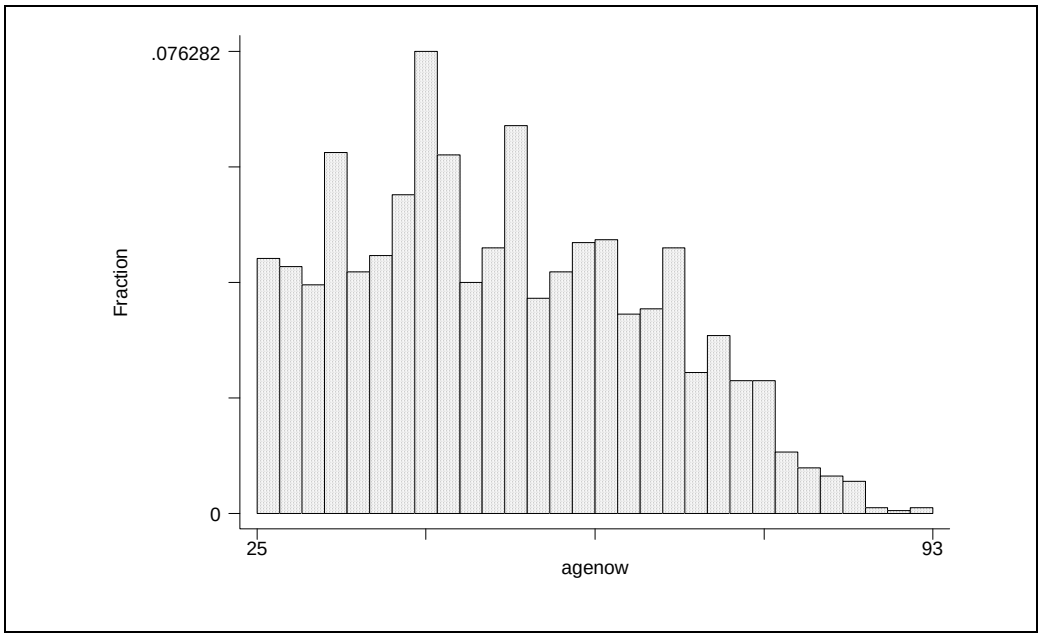


Figure 5-3 Distribution of respondent ages in estimating sample

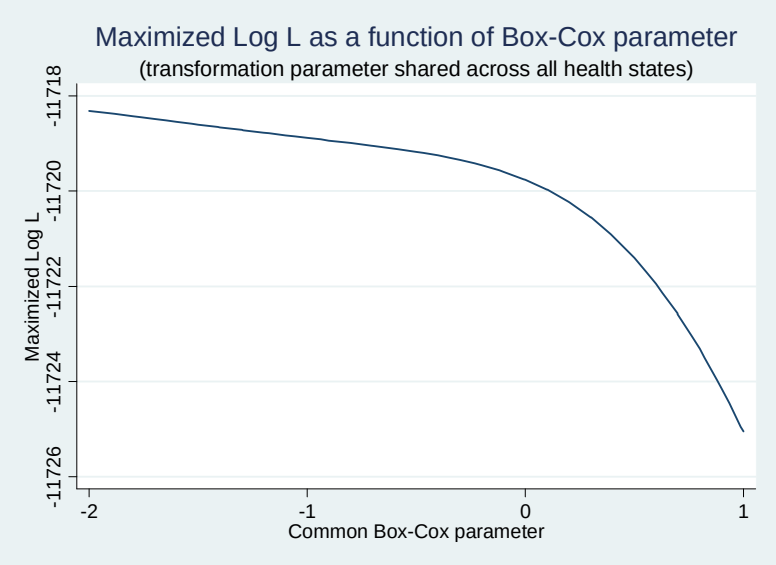


Figure 5-4 Log L as a function of health state duration transformation Rationale for using shifted log transformation of each discounted prospective health state duration: consequences of line search across a common Box-Cox transformation parameter for each of the three shifted adverse health state durations, with zero value implying logarithmic transform. Preliminary specification with selection correction but without scenario adjustments

You may have chosen Program A, Program B, or neither. Regardless of your choice, we would like to know when, over your lifetime, you think you would first need and benefit from the two programs (if at all).

Your answers below may depend upon the illness or injury in question, as well as your current age, health and family history.

Around when do you think you would begin to value highly the risk reduction benefits of each program?

Select one answer from each column in the grid

	Program A to reduce my chance of diabetes	Program B to reduce my chance of heart attack
For me, benefits would start:		
Immediately	☐	☐
1-5 years from now	☐	☐
6-10 years from now	☐	☐
11-20 years from now	☐	☐
21-30 years from now	☐	☐
31 or more years from now	☐	☐
Never (Program would not benefit me)	☐	☐

Figure 5-5 Example of debriefing question for scenario adjustment
The "Never (Program would not benefit me)" response is very unambiguous; however, subjective estimates of when the program would "begin to benefit" the individual do not map as cleanly relative to the illness profiles described in each choice question.

We cannot perfectly predict how long we will live. But based on our health and family history, most of us have some idea about how long we might live.

Until what age do you expect to live? Please check your best guess.

Select one answer only

◉ 54	○ 65	○ 76	○ 87	○ 97
○ 55	○ 66	○ 77	○ 88	○ 98
○ 56	○ 67	○ 78	○ 89	○ 99
○ 57	○ 68	○ 79	○ 90	○ 100
○ 58	○ 69	○ 80	○ 91	○ 101
○ 59	○ 70	○ 81	○ 92	○ 102
○ 60	○ 71	○ 82	○ 93	○ 103
○ 61	○ 72	○ 83	○ 94	○ 104
○ 62	○ 73	○ 84	○ 95	○ 105
○ 63	○ 74	○ 85	○ 96	○ More than 105
○ 64	○ 75	○ 86		

Figure 5-6 Debriefing question about life expectancy

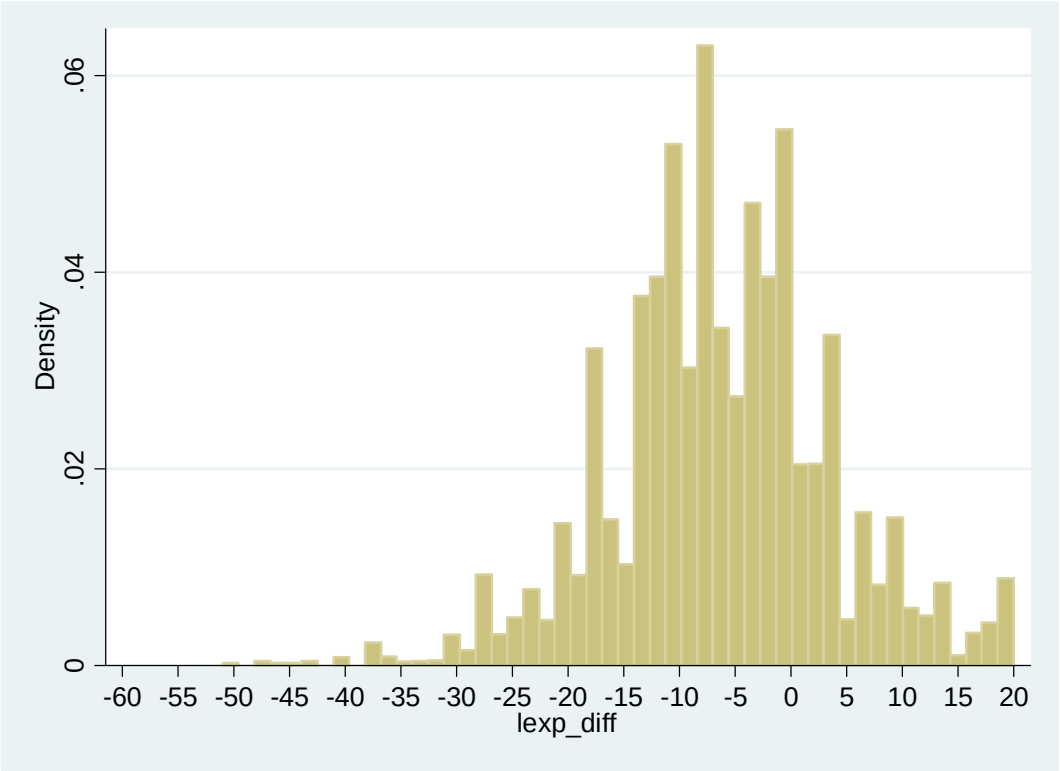


Figure 5-7 Histogram: Subjective overestimates of life expectancy
On average, subjective life expectancy is 6.28 years less than the age we used in the choice scenarios to describe the different illness profiles. We added eight years to avoid respondents' rejections of actuarial life expectancy, but it seems that on average, we needed to add only about two years. However, life expectancies more than 25 years less than we told people seem questionable, unless the individual already has a terminal illness. Thus we use the logarithms of the absolute values of these departures in our scenario adjustment methods. This decreases, without omitting, the influence of the large negative outliers.

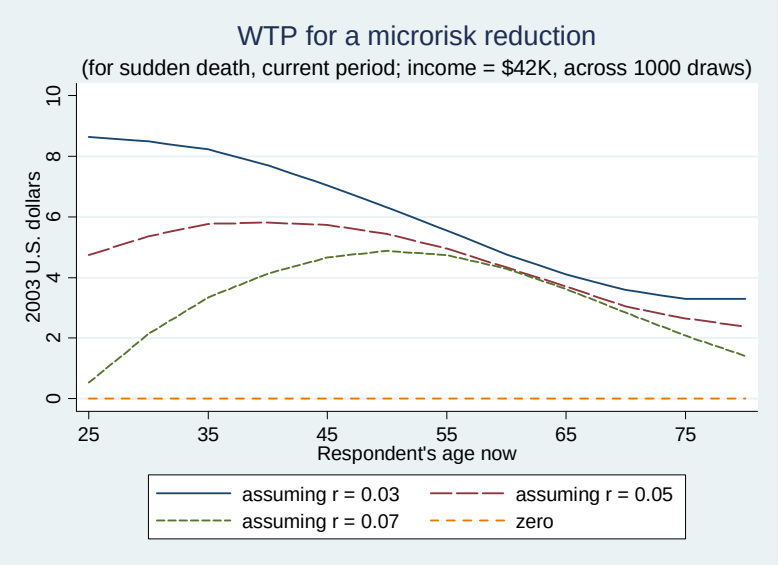


Figure 5-8 WTP, sudden death now, by discount rate
WTP for a microrisk reduction for sudden death now, as a function of respondent age now, for three different discount rate assumptions; Based on preliminary specification with three exclusion criteria and no scenario adjustment/rejection controls.

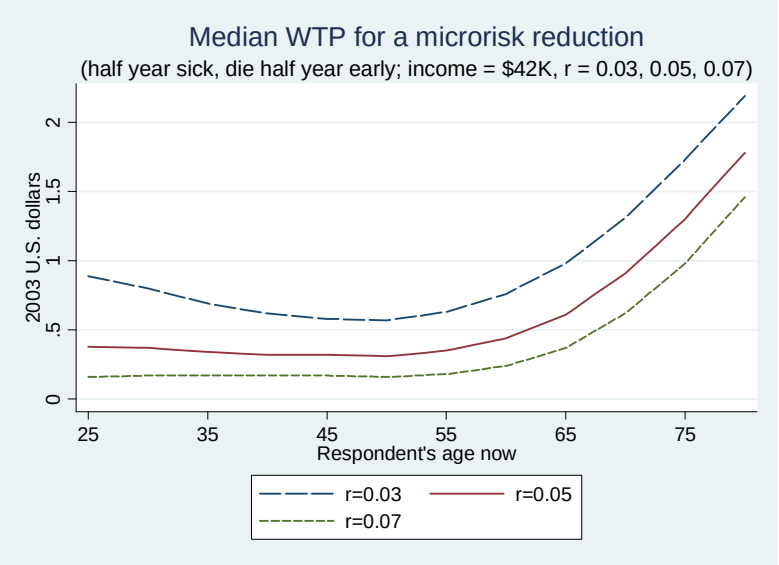


Figure 5-9 WTP, half-year sick, die half-year early, by discount rate
WTP for a microrisk reduction for six month reduction in life expectancy, preceded by six months of major illness, as a function of age now, for three different discount rate assumptions; Based on preliminary specification with three exclusion criteria and no scenario adjustment/rejection controls.

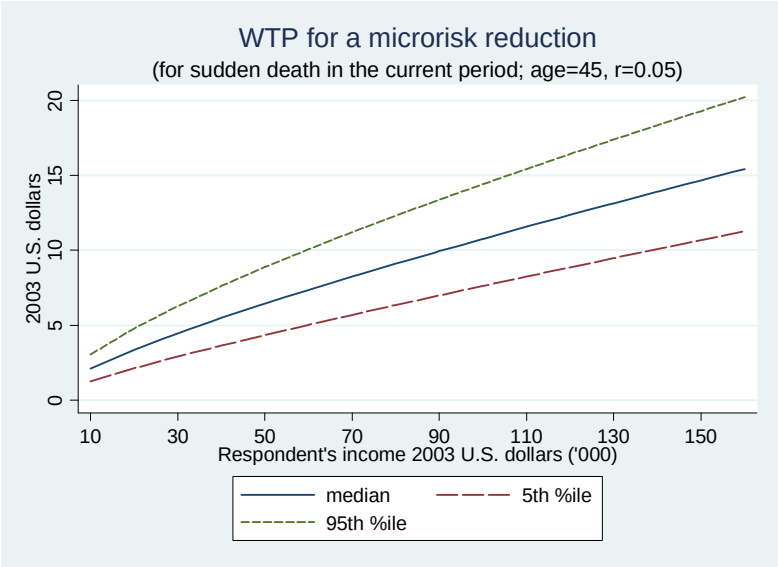


Figure 5-10 WTP, sudden death now, by income level
WTP for 1/1,000,000 reduction in risk of sudden death in the current period, as a function of respondent household income now in $'000, for a 45-year-old; Based on preliminary specification with three exclusion criteria and no scenario adjustment/rejection controls.

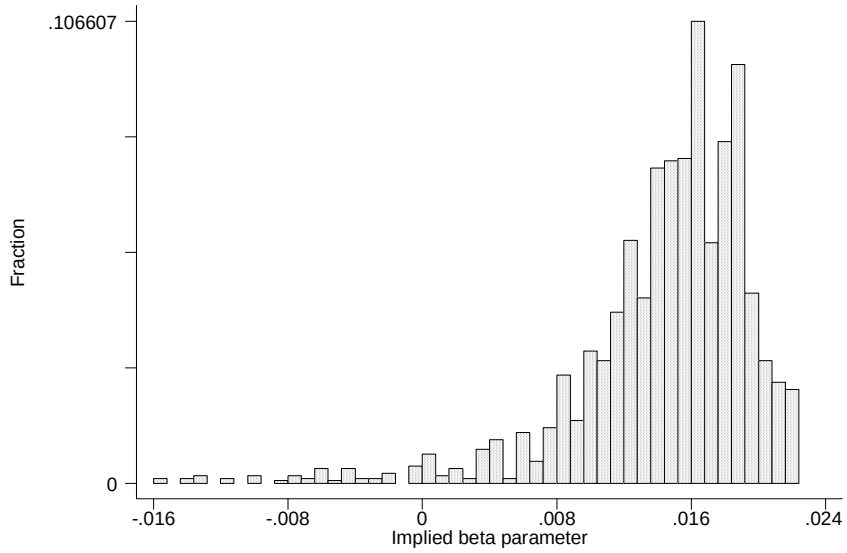


Figure 5-11 MU(Y) by "room to improve" and "difficulty of lifestyle changes" Preliminary specification with selection correction but without scenario adjustment

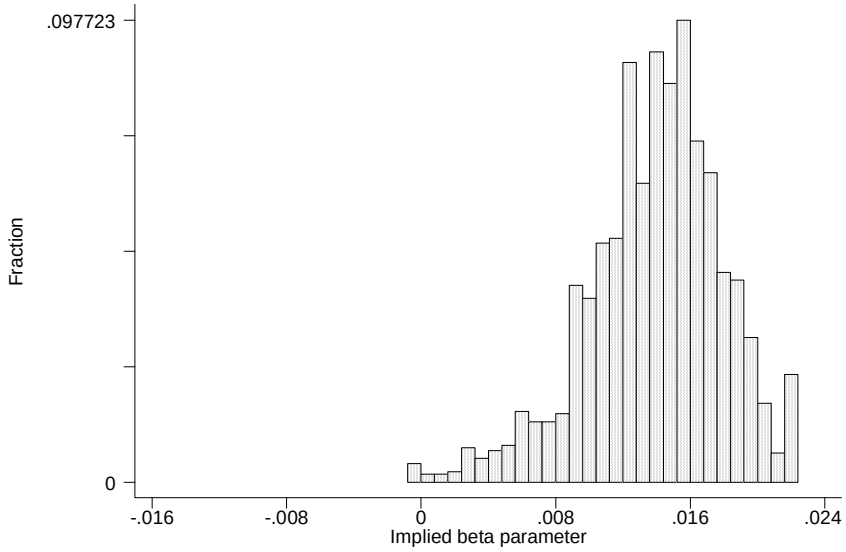


Figure 5-12 MU(Y) as a function of "difficulty of lifestyle changes" only Preliminary specification with selection correction but without scenario adjustment

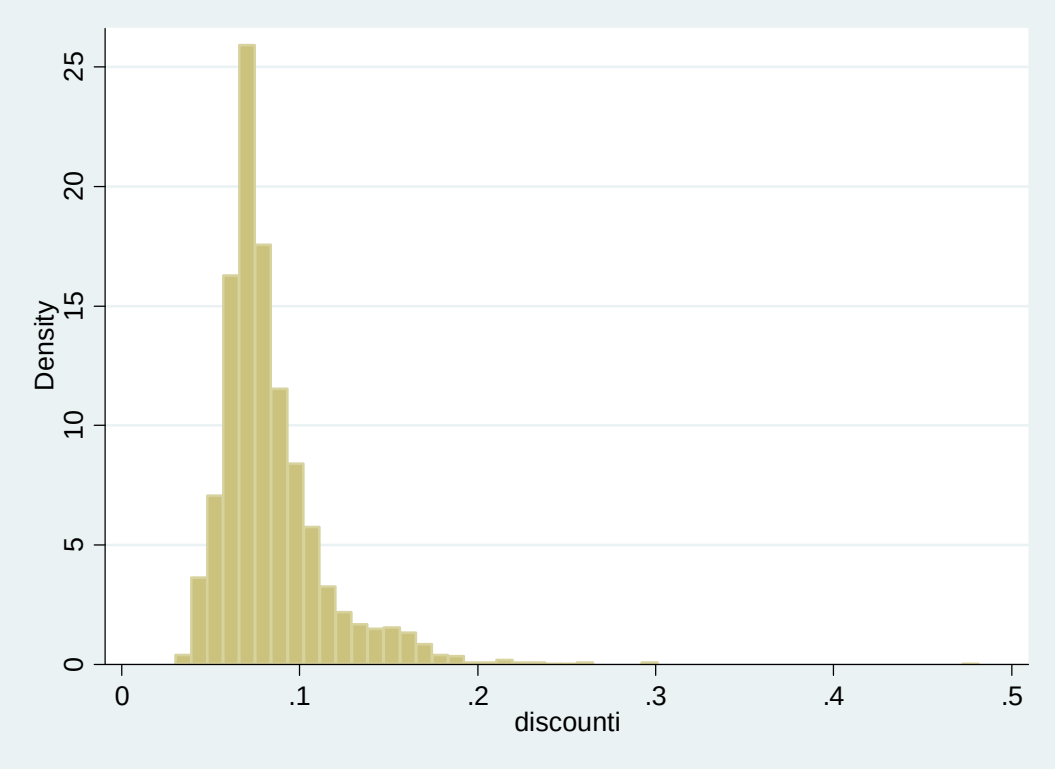


Figure 5-13 Distribution of calculated individual discount rates
Rates are based on discounting model for the "public choices" survey as derived in Bosworth et al. (2011), applied to the corresponding variables for respondents to the "private choices" survey.

6 Inventory of Research Papers using these Data

Our survey was designed to make it possible to analyze a wide range of research questions. One important advantage of stated preference conjoint choice experiments is that the random assignment of attribute levels across alternatives and choice sets precludes any correlation of these attributes with the characteristics of the respondent (beyond the unavoidable age and gender characteristics which we use to define the range of possible illness profiles for each individual). This absence of correlations means that it is possible to analyze the effects of subsets of program attributes without worrying about omitted variables bias that is might produce if one were to use real choice data. This advantage means that we can address different types of program attributes in different papers, rather than needing to control for all possible attributes in any single specification.

All of the papers which rely on the data from this survey refer to the "flagship" or "main" paper that lays out the basic model.

6.1 "Flagship" or "Main" paper:

Cameron and DeShazo (2011) "Demand for health risk reductions," manuscript

Abstract: A choice model based on utility in each of a sequence of prospective future health states permits us to generalize the concept of the Value of Statistical Life (VSL). Our representative national survey asks individuals to choose between costly risk-reducing programs and the status quo in randomized stated choice scenarios. We estimate separate marginal utilities for discounted net income and avoided illness years, post-illness years, and lost life-years. Our estimates permit calculation of overall willingness to pay to reduce risks for a wide variety of different prospective illness profiles. These can be benchmarked against the VSL as a special case.

6.2 "Kids" paper

Cameron et al. (2010a) "The effect of children on adult demands for health risk reductions," *Journal of Health Economics* 29(3) 364-376.

Abstract: We examine patterns in adults' willingness to pay for health-risk reductions. We allow both their marginal utilities of income and their marginal disutilities from health risks to vary systematically with the structures of their households. Demand by adults for programs which reduce their own health risks is found to be influenced by (1) their parenthood status, (2) the numbers of children in different age brackets currently in their households, (3) the ages of the adults themselves, (4) the latency period before they would fall ill, and (5) whether there will still be children in the household at that time. For younger adults, willingness to pay by parents is greater than for non-parents, and increases with each additional young child. For middle-aged adults, willingness to pay for corresponding risk reductions falls when teenagers are present and falls further with each additional teenager in the household.

NOTES: This paper inadvertently omits to mention that WTP simulations are based on an original risk of 0.004 and a risk reduction of 0.003 (leaving a revised risk of 0.001). As demonstrated in is supplementary document, altering the original risk appears to have a relatively modest effect on the estimates of average WTP over the risk change for a one-microrisk reduction. Scenario adjustment variables used in this paper are the "never benefit" indicator and the "overestimate of the latency," but not the "overestimate of life expectancy." The function $f(\cdot)$ in this paper is assumed to be quadratic in net income.

6.3 "Canada" paper

Cameron et al. (2010c) "Demand for health risk reductions: A cross-national comparison between the U.S. and Canada," *Journal of Risk and Uncertainty* 41(3) 245-273.

Abstract: Using a large stated preference survey conducted across the U.S. and Canada, we assess differences in individual willingness to pay (*WTP*) for health risk reductions between the two countries. Our utility-theoretic choice model allows for systematically varying marginal utilities for avoided future time in different adverse health states (illness-years, recovered/remission years, and lost life-years). We find significant differences between Canadian and U.S. preferences. *WTP* also differs systematically with age, gender, education, and marital status, as well as a number of attitudinal and subjective health-perception variables. Age profiles for *WTP* are markedly different across the two countries. Canadians tend to display flatter age profiles, with peak *WTP* realized at older ages.

NOTES: This paper assumes that the function $f(\cdot)$ is the square-root function, approximating a Box-Cox transformation parameter of 0.5, compared to the 0.45 used in some of our other papers. The impact of this adjustment is modest. The scenario adjustment variables employed in this paper include the "never benefit" indicator, and both the overestimate of the latency and the overestimate of life expectancy as single continuous variables which enter linearly.

6.4 "Diseases" paper

Cameron et al. (2011) "Willingness to pay for health risk reductions: differences by type of illness," manuscript under review

Abstract: Our research identifies large systematic differences, by type of illness, in individual willingness to pay (WTP) to reduce the risk of the major health threats. These include 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. Our estimates take the form of individuals' WTP to reduce the risk of experiencing specific illness profiles (i.e. the different patterns of sick-years, recovered/ remission-years and/or lost life-years associated with each illness). Our results suggest that analyses which constrain the marginal utility parameters for different health states to be the same across all illnesses are too restrictive, causing the loss of valuable information for benefit-cost analyses of health, environment and safety policies. We also find that the rank ordering of

private willingness to pay for illness-specific risk reductions is highly correlated with public spending patterns by government agencies.

6.5 "Scenario adjustment" paper

Cameron et al. (2010b) "Scenario adjustment in stated preference research," *Journal of Choice Modelling 4(1)*, 9-43.

Abstract: Poorly designed stated preference (SP) studies are subject to a number of well-known biases, but many of these biases can be minimized when they are anticipated ex ante and accommodated in the study's design or during data analysis. We identify another source of potential bias, which we call "scenario adjustment," where respondents assume that the substantive alternative(s) in an SP choice set, in their own particular case, will be different from what the survey instrument describes. We use an existing survey, developed to ascertain willingness to pay for private health-risk reduction programs, to demonstrate a strategy to control and correct for scenario adjustment in the estimation of willingness to pay. This strategy involves data from carefully worded follow-up questions, and ex post econometric controls, for each respondent's subjective departures from the intended choice scenario. Our research has important implications for the design of future SP surveys.

> *NOTES: This paper makes the methodological case for attention to the possibility of scenario adjustments that fall short of outright scenario rejection. It demonstrates the effects of including and excluding the "never benefit" indicator and the "overestimate of the latency" variable as an illustration of the potential influence of corrections of this type.*

6.6 "Attention to attributes" paper

Cameron and DeShazo (2010b) "Differential attention to attributes in utility-theoretic choice models," *Journal of Choice Modelling, 3(3)* 73-115.

Abstract: We show in a theoretical model that the benefit from additional attention to the marginal attribute within a choice set depends upon the expected utility loss from making a suboptimal choice if it is ignored. Guided by this analysis, we then develop a very general and practical empirical method to measure an individual's propensity to attend to attributes. As a proof of concept, we offer an empirical example of our method using a conjoint analysis of demand for programs to reduce health risks. Our results suggest that respondents differentially allocate attention across attributes as a function of the mix of attribute levels in a choice set. This behaviour can cause researchers who fail to model attention allocation to estimate incorrectly the marginal utilities derived from selected attributes. This illustrative example is a first attempt to implement an attention-corrected choice model with a sample of field data from a conjoint choice experiment.

> *NOTES: These data are used to illustrate another methodological point that researchers may wish to consider in choice-based research. The theoretical portion of this paper predates our survey, which turned out to provide a useful set of empirical data to use as an example.*

6.7 "Choice difficulty" paper

Duquette et al. (2010) "Subjective choice difficulty in stated choice tasks," manuscript

Abstract: Objective dimensions of choice set complexity (often measured only in attribute space) have been used as a proxies for choice difficulty, and these proxies have been used empirically to shift the scale of the error term or the slope coefficients in choice models. However, the full scope of "choice difficulty" is not usually observable by the researcher, since choice difficulty may also depend upon the characteristics of the individual who is trying to make the choice. In our stated preference survey, respondents are asked directly to rate the subjective difficulty of each of their choices. We use this difficulty rating to assess how well the customary reduced-form complexity proxies are likely to capture this aspect of subjects' interactions with choice tasks. Common measures do not fully explain subjective choice difficulty, which depends on the interplay among objective attribute-space complexity, the similarity of alternatives in utility space, a variety of respondent characteristics and cognitive resource constraints. Subjective choice difficulty also appears to have systematic effects on estimated preference parameters and implied WTP estimates.

6.8 "Age" paper

DeShazo and Cameron (2006b) "Two types of age effects in the demand for reductions in mortality risks with differing latencies," manuscript

Abstract: We develop and test an empirical model of individuals' intertemporal demands for programs to mitigate health risks over the remaining years of their lives. We estimate this model using data from an innovative national survey of demand for preventative health care. We find qualified support for the Erhlich (2000) life-cycle model, which predicts that individuals expect to derive increasing marginal utility from reducing health risks that come to bear later in their lives. However, we also find that as individuals age, there appears to be a systematic downward shift in their anticipated schedule of marginal utility for risk reduction at future ages. Our model improves upon earlier work by differentiating between the respondent's current age and the future ages at which they would experience adverse health states. Using estimated demand schedules specific to an individual's current age, we demonstrate the calculation of values for risk mitigation programs that reduce the probabilities of specified time profiles of adverse future health states involving various latency periods.

> *NOTES: Paper has been essentially sidelined for a number of years, with renewed effort awaiting the final disposition of the "main" paper from this study.*

6.9 "Comorbidity" paper

DeShazo and Cameron (2006a) "The effect of health status on willingness to pay for morbidity and mortality risk reductions," manuscript.

Abstract: Both actual and expected morbidity systematically affect individuals' demands for both life-saving policies and preventative health care. Using a large general-population sample,

we estimate a utility-theoretic model of consumer preferences across risk reduction programs targeted at a wide variety of major health threats with differing illness profiles. Individuals' demands for programs targeting a particular illness are higher when there is a history of that illness and when subjective risks are higher. A history of other illnesses and greater other-illness subjective risks decrease demand. These comorbidity effects operate through the marginal utilities of both (i) adverse health states and (ii) income.

NOTES: Paper has been essentially sidelined for a number of years, with renewed effort awaiting the final disposition of the "main" paper from this study.

7 Full Disclosure of Peer Review History

The peer review process is crucial for any research that might be used to support public policy-making. If it can be avoided, of course, most authors would probably prefer to suppress the fact that their paper was rejected by other journals before its eventual publication. In this case however, given the very high stakes associated with the choices made by government agencies for the values of environmental, health, and safety risk reductions, we take the unusual step of revealing the very helpful comments and concerns voiced by various editors and reviewers who have considered previous versions of our main paper. This dialogue reveals the extent of the cumulative scrutiny of our work. We include the verbatim comments of our editors and referees and our responses and rebuttals to these comments. We omit only those comments that concern trivial points (such as typographical errors).

Importantly, these comments refer to obsolete archival versions of the specific manuscripts upon which each set of comments was based. We have certainly learned where our ideas and procedures need to be explained better, and we have incorporated these insights into the latest version of our main paper. In some cases, our responses to the referees were invited by the editor of the journal in question. In other cases, there was a summary rejection based upon the opinions of the reviewers (or even just the editor) with no invitation for us to respond. Despite these dead ends, we took seriously the comments of every reviewer and we either rebut their mistaken impressions about the research or explain in more detail how we accommodate their concerns in subsequent revisions of the paper.

Our main paper in its current form owes much to these earlier editors and reviewers. From them, we have learned a lot about how much infrastructure must be provided before an uninitiated reader can digest the work that is described in the papers associated with the project. The reactions of many readers—in particular, those at the two extremes of the spectrum (i.e. those who are almost completely unfamiliar with non-market valuation of reductions in risk to life and health, and the most expert specialized referees)—have led us to refine the main paper and also to resort to this lengthy appendix-like document to address both the broad-brush and detailed concerns which have been raised. Standard overall journal-article length limits of about forty total manuscript pages do not permit us to cover all of these points in each paper in the series.

It is worth noting that economists who depend upon "revealed preference" data provided by government agencies are typically very trusting that someone else has worried about the representativeness of the sample and the quality of the data (whether or not that trust is warranted). For a study like this one, however, which employs original researcher-collected data based on stated-preferences, many issues of data quality must be addressed in minute detail before readers can be confident in the usefulness of the data for policy-related analysis. This is as it should be, of course. This supplementary document attempts to fill that need, with some economies of scale across the variety of papers associated with this survey.

7.1 Early submission to the *Journal of Political Economy*

First, we appreciate the feedback from the *Journal of Political Economy*, where we sent the original very succinct version of this paper in August of 2004. Our paper was assigned to an editor whose own fields were labor markets and macroeconomics. He procured one review and

rejected the paper. The reviewer and the editor had four major concerns: (1) that the paper had nothing to do with financial options (we were not clear enough that we use "option price" in the sense of Graham (1981), rather than in the financial sense); (2) that it was absurd to conclude that an individual would pay the equivalent of 56 years of income avoid death (we did not adequately tutor our audience about the definition of the Value of a Statistical Life and the fact that it is not the same thing as an individual's willingness to pay for a 100% risk reduction); (3) that we failed to control for all of the illness profile attributes that differ across alternatives and therefore our estimates were biased (the attributes in a conjoint choice "experiment" are randomized so that excluded variables are uncorrelated with included variables by design, so omitted variables bias is not a significant concern), and (4) that our method does not estimate the cost of illness (when our point is that we are *not* trying to measure the cost of illness, but instead to measure subjective benefits, which is the preferred approach to welfare estimation).

The misconceptions generated by this early version of the paper thus included a number of problems stemming from the editor's and referee's lack of familiarity with the concept of the "value of a statistical life" (despite the fact that footnote 1 in the paper explained it). This degree of confusion about the VSL, even among economists, not to mention among members of the general public, was the impetus for Cameron (2010). We were also hampered by our failure to emphasize that the fact that the term "option price," as it is used in environmental economics, has a different interpretation than is used in the finance literature.

Based upon the referee's comments and his own reading of the paper, the editor of the *JPE* decided to reject our paper without permitting us to respond to those comments. We did take the unusual step of responding to them anyway, in an attempt to clear up some significant misunderstandings about our work that led the editor and the referee to be unimpressed by it.

Since our rebuttal was not communicated to our referee, and the editor did not acknowledge the receipt of our email, we include our replies here, where we have added a few more details to bring our replies up to date to reflect the current version of our main paper. The anonymous referee remains out there somewhere and may have been left with the erroneous impression that their most significant criticisms were valid. We certainly revised the paper substantially in light of these comments.

7.1.1 *JPE* Editor's and referees' comments, and our replies

Editor's comments:

…I have received one detailed referee report and have read your paper closely myself. The referee recommends in a cover letter that "a substantially revised (and shorter) version might be of interest to the *Journal of Public Economics.*" I agree with both parts of that assessment: (i) the paper needs to be substantially revised and much more focused and (ii) even then, it is unlikely to find a place at the *JPE.* It is better-suited to a journal specializing in health economics.

I found the way you describe both the model and the survey very confusing. I'll start with the modcl. You describe the theory as a "structural option price model." I don't see anything that is normally associated with options in the paper, e.g., a decision on when to exercise the option. There is nothing wrong with that, but your language makes it hard to understand what is happening.

RESPONSE: We should have been more explicit that Graham's 1981 AER concept of "option price" has nothing to do with "options" as the term is used in the finance literature. It stems instead from the consumer theory literature on benefit-cost analysis of public programs under uncertainty, which is the application for the empirical results we derive.

As far as I understand, you assume expected indirect utility is a quadratic function of expected income after health expenditures and of the expected amount of time spent in three different health states. The latter is a function of health expenditures, and income also depends on whether a particular health state is realized. An individual chooses a health program to maximize her indirect utility. You ask how much an individual would be willing to pay per year for a particular health program, or alternatively convert this in to a utility-equivalent one-time expenditure.

RESPONSE: These particular assumptions are only made for the simplified specification we use to develop the intuition behind the model. The actual estimating specification for the model is more complex, since the data dictate systematic variation of marginal utilities with respect to respondents' ages and nonlinearity of utility in discounted lost life-years, as well as an interaction term between sick-years and lost life-years. To conform with the usual measure of costs for public policies and regulations that reduce health risks, our WTP measures are *per year* (for consumers of different ages and incomes, who face different types of prospective illness profiles).

The referee claims that your model could be reduced from 18 to 10 pages. I'd say it could be about three pages long and still make the same point, primarily equation (6).

RESPONSE: Lack of familiarity with VSLs probably explains why you and the referee were of the opinion that the model could be three pages long and end with equation (6), and why you thought that the material between equation (6) and equation (13) was a bunch of superfluous "examples." Actually, that stuff is essential to the process of getting from a simple annual WTP measure to something that can be benchmarked against the $6 million figure used by the US EPA in benefit-cost analyses of all its major environmental regulations, such as the Clean Air Act, for example. (The Department of Transportation employs a number closer to $1 million.) We elected to use a simplified model to outline the algebra. It is necessary to show the basic steps in the process. We thought it would have been unnecessarily complicated to show these constructs in terms of the actual working model that involves risk-averse, rather than risk-neutral, preferences that are also heterogeneous in age and a flexible second-order translog-type approximation to actual preferences.

The paper's primary contribution is your survey. To be frank, I do not understand why you set up the survey the way you did. As the referee notes, "to be ill with cancer. diabetes, stroke, and Alzheimers [sic] (for instance) arc very different experiences." It is unclear why you would expect to obtain a single measure for the annual cost of illness. Moreover, you admit that survey respondents likely carried their own prejudices into the survey, incorporating these into the costs of different illnesses. This undoubtedly biases your estimates of the cost of illness, recovery, and death. Footnote 20 needs to be addressed directly in any publishable version of this paper.

Probably you need to estimate the cost of each illness separately, which goes against the entire spirit of your competing-risks framework.

> RESPONSE: We feel that the primary contribution is our conceptualization of illness profiles as mixes of time periods in different health states, and the design of a survey that allows empirical characterization of preferences with respect to reductions in the risk of suffering these illness profiles. We are not estimating the "annual cost of illness" as you suggest. There are "Cost-of-Illness" calculations that are sometimes used as placeholders when no utility-theoretic demand information for health-risk reductions (like ours) is available. This paper is about ex ante perceived benefits, not ex post costs.

One important methodological point: if I understand Table 3, it seems dangerously close to curve-fitting. I am very skeptical of the improvements you found in the log-likelihood function.

> RESPONSE: There are familiar theoretical reasons to entertain specifications where utility is diminishing in the quantities of the goods in the consumption bundle. Constant marginal utilities and constant marginal rates of substitution between goods should generally not be assumed unless one fails to reject this assumption, based upon a more flexible model that allows for more generality. Translog functional forms are very popular as local approximations to arbitrary preferences.

My last point concerns the bottom line, 4 and 5. I find the results implausible. A 45-year-old individual with (pre-tax?) income of $42,000 is willing to spend 56 years' income to avoid one year of sickness followed by full recovery. No one cares much, at least in a statistical sense, between immediate death and five years of sickness followed by death. One can accept these results at face value or one can conclude that when confronted with screens like that in Appendix A, people are either confused or select options with little regard to the true implications. I put high probability weight on the latter possibility. This suggests that you need to be much more cautious in interpreting your results or you need to provide much more evidence that convinces the reader that the survey instrument elicits accurate responses.

> RESPONSE: A WTP of "56 years' income" for a health risk reduction would certainly be ridiculous. But a VSL is not a WTP. The VSL implied by an individual's choices is not constrained by their income. If 1 million people are each willing to pay $6 for a 1-in-1-million risk reduction, the corresponding VSL is $6 million dollars. VSLs are WTP numbers for very small risk reductions, scaled up to a 1.00 risk change (but with no pretense that a 1.00 risk change is ever at stake).
>
> We view our results as greatly reassuring that people gave our choice scenarios due consideration and responded fairly thoughtfully. People's choices, in combination with our model, yield WTP predictions (for the special case of our more-general model that corresponds to the VSL) that align nicely with most of the existing estimates gleaned from a wide variety of other studies. We take this correspondence as "evidence that convinces the reader that the survey instrument elicits accurate responses." In our revisions to the paper, we will emphasize this correspondence more prominently.

Referee's comments

Theory

The main issue I have with the paper is conceptual. To be ill with cancer, diabetes, stroke, and Alzheimers (for instance) are very different experiences and probably should not be treated as identical "illness states". Ditto for the "recovered" state. For example, those with diabetes may be at greater risk of cardiovascular disease, kidney disease, blindness etc but according to the paper, the period before these complications of diabetes occur is equivalent in utility terms to the case where they occur. This is problematic given that the model claims to be based on EU. The diabetes problem also highlights a second major problem, in that each disease category is composed of many separate health states that will also differ in utility terms from each other. Any VSI using broad disease categories cannot give anything better than a rough estimate of benefit; there is no guarantee that the VSI method provides a better solution than the inconsistent piecewise approach that the paper seeks to replace.

> RESPONSE: Many existing estimates of WTP to reduce health risks focus solely on the risk of sudden death in the current period from fatal on-the-job accidents. Our survey encourages people to think about a wide range of health threats with different features, including many with long latency periods before symptoms appear and/or long periods of morbidity before death. There is a great deal of heterogeneity in the variety of disease names we associate with the illness attributes that are arguably the proximate sources of disutility: moderate versus severe pain and disability, hospitalization, surgery, latency, total sick-time, any recovered/remission time, and lost life-years. These common attributes are shared across all of our illness profiles, in different combinations of levels or quantities. It is also possible that the identical disease profile can be associate with two different disease names.
>
> Fortunately, it doesn't matter that we leave out the illness names, because they are randomized. Eligible illness profiles depend only upon the respondent's age and gender, arguably the most exogenously determining personal attributes in our usual set of candidates. Omitted variables that are uncorrelated with included variables do not produce omitted variables bias. Even the extra noise still leaves us with hugely significant parameter estimates. Many empirical researchers are used to non-experimental data and have every expectation that leaving out something that is obviously important will produce omitted variables bias. However, the beauty of stated-preference methods is that the researcher gets to design the matrix of "regressors." Inclusion of the illness names reduces the error variance and lets you look for systematic effects of these names, but excluding them causes little mischief, fortunately.

The method in the paper differs from standard VSL in that it attempts to take a multi-period view. This takes approximately 18 pages of the text but should be reduced to under 10 pages by the judicious removal of matter that's either trivial or not required by the subsequent empirical work. Some suggestions to achieve this are:

> RESPONSE: We DO take a multi-period view, in terms of prospective future illness profiles involving many different (arbitrary) patterns of health states. Breaking away from the tradition of focusing mostly on a single profile, consisting of sudden death in the current period, or just one or two specific profiles, is our major contribution.

• The authors might consider revising the text to use vector notation, which can summarise much of the algebra far more quickly.

RESPONSE: The nonlinear form of the actual estimating model does not lend itself to much simplification through vector notation (as opposed to the simplified linear version used to demonstrate the approach).

• The paper makes a valid case early in the piece that three of the four gamma coefficients it defines should be zero and uses these values within the empirical work. As it stands, little justifies subsequent equations including these coefficients.

RESPONSE: Revisions to the paper, and sensitivity analyses, explore the consequences of different assumptions about each of the gamma coefficients, also at the request of subsequent referees. The conditions used for simulation of WTP may differ from those for the choice tasks and the estimation. For example, private goods are used in estimation, whereas public goods may be relevant when WTP is being calculated. The gamma parameters are needed to preserve our ability to simulate WTP in circumstances where individuals probably make different assumptions about their future income and future program costs than they do in the context of our choice exercises.

• There is a lot of basic algebraic manipulation. For example, the authors repeat a lengthy equation in a rearranged form on Pages 6 and 7. Any reader with a basic grasp of mathematics - and I'm sure the JPE readers easily manage this - are able to verify that the equation is quadratic in cost easily.

RESPONSE: To make life easier for our referees, we consciously erred on the side of providing more detail in the derivations of our formulas. We are certainly grateful for the editorial suggestions and concur that the paper can be shortened easily if the consensus is that readers do not need to see the derivations in such detail. In that archival version of the paper, the quadratic-in-cost nature of preferences was not expected to be nearly as challenging as the final formulas used to simulate the sampling distribution of WTP under different circumstances, based on the joint distribution of the estimated parameters.

• Due to the lengthy nature of the equations used, the paper also resorts to a simplified case in order to give examples. These examples add very little to the paper, aren't used in practice, and detract from the flow what should be a much quicker introduction to the empirical method.

RESPONSE: Perhaps our strategy was in error. The alternative was to jump right into the nonlinear-in-income and translog-in-health states specification where the coefficients also vary with age and age-squared. We felt it would be unnecessarily complicated, in that case, to show the steps in getting from the conditional logit parameter estimates to the resulting WTP formula under counterfactual conditions. So we decided to outline the general strategy in a simplified case and appeal to the reader to appreciate that the estimating specification actually used would be more complex. It probably wouldn't be

wise to dispense with the simplified exposition altogether. It is also important to realize that these are not "examples" that add very little to the paper. Instead, they form the core of our exposition of the steps for getting to the WTP amounts of interest.

Methodology

As the authors identify nine important attributes of interest across the disease areas and use four of them (plus disease label) in the empirical work. Given that there are at least five important unlisted attributes that may be inferred from the disease label alone, it seems very unlikely that only the listed attributes will be considered. However, this is precisely what conjoint analysis requires. With the current methodology, the results found are theoretically unsound. Whilst the authors wish to reserve results with illness-specific dummies to a subsequent paper, I doubt this initial paper is publishable without them (at least in a good journal in either economics or health economics).

RESPONSE: We suspect that this referee did not catch the fact that the disease attributes describe "stylized" disease, not actual diseases. The illness names are designed to be essentially orthogonal to the disease attributes, except in a few cases where an illness profile for a particular disease would be implausible. For example, we did not allow for sudden death or recovery from diabetes or Alzheimer's disease, although there were many different latency periods, sick-years, and lost life-years across all of the choice contexts that included these diseases. Our supplementary documentation provides details of the experimental design of the different mixes of illness profile attributes used in our study.

Suppression of the illness-specific dummies means that the WTP amounts we derive for our different illness profiles represent an "average" across all of the twelve different types of disease labels employed in our survey. However, these illnesses and injuries are heterogeneous in terms of their time profiles for different health states, so WTP is heterogeneous along these dimensions as well. WTP also depends fundamentally on income and on age at the time the program choices are being made. This makes our model, and the types of demand measures it can generate, far more general that the conventional "VSL" which typically concerns WTP to reduce just one illness profile: the risk of sudden death in the current period. These ordinary VSL numbers have been used for decades to represent people's WTP to reduce all kinds of health risks, and in many cases those health risks are rather remote in their time-profiles compared to "sudden death in the current period."

With the very long outline of the algebra, several aspects of the empirical method aren't adequately covered. For example, how many questions did each individual face? If this figure is too large, then the empirical literature would suggest fatigue will have degraded the quality of results obtained. Further, the paper neglects to provide an indication of how prevalent nontrading behaviour appears to be, which is uncommon in my experience.

RESPONSE: This information appeared on the first line of page 23 in the submitted paper, in the description of the survey instrument. I guess we should mention it more than

once. Our supplementary documentation now reviews the structure of the survey in much greater detail.

Non-trading behavior (i.e. choice of the "No Program") alternative is balanced quite well with trading behavior. Based on our pre-tests, we adjusted the distribution of program costs to ensure that there would be variability in the relative attractiveness of each alternative, including the "no program" alternative. In our estimating sample that includes 1801 individuals and 7520 distinct choices, respondents chose the program on the left of the pair 34.89% of the time; they chose the program on the right of the pair 34.24% of the time, and they chose the "no program" option 30.86% of the time. Keep in mind that we rejected choice sets in cases where the respondent chose the "no program" alternative, but gave as their ONLY reason for doing so the explanation that they did not believe the program would work. These instances display outright scenario rejection by the respondent, so we argue that we cannot expect choices under these circumstances to be consistent with our model.

Removing those that selected "neither program" because they did not believe a program would work may be quite understandable in the presence of a threshold effect. Threshold effects are not mentioned.

RESPONSE: We assume that by "threshold effect," this referee is referring to a status quo effect, where respondents will stick with the status quo unless some threshold of prospective utility is available from one of the risk-reduction program alternatives. We contend that removal of choices made under circumstances where respondents admit rejection of the choice scenario should be done a priori. The idea behind threshold effects is that subjects accept the attribute levels and the story behind the choice set, yet there is some non-specific "lump" of utility associated with making no change and sticking with the status quo. Our current models now include a status quo effect, but this is something quite different than an a priori criterion for exclusion of certain choices from the estimating data.

Results

I have some concerns as to the policy relevance of the results.

- We are not told how many people have predicted negative VSI values. From the text it appears that these are partially determined by age and income, and so older poorer people may face negative VSI values for many of the health-improving treatments they face. This may have important equity considerations that aren't explored.

RESPONSE: Our data provide no information at all about negative WTP for any of these programs. The worst that anyone could do on a choice occasion was to choose the "neither program" alternative. When we draw 1000 sets of utility parameters from the joint distribution of the estimated parameters, there are some simulation scenarios for which some sets of parameter values produce a negative value for calculated WTP. However, we invoke a Tobit-like interpretation in these cases and interpret these negative values as zero. While there may in fact be latent negative WTP in a few cases, there is no way in our choice scenarios (or in most real markets) to exercise this negative WTP—

namely where the individual would have to be paid to accept a reduction in the risk to their health, ceteris paribus.

"Older poorer people" may indeed be willing to pay little or nothing to reduce their risks of certain health threats, perhaps because the latency of that threat is likely to exceed their remaining lifespan, or they expect to die from something else first. It is of course true that when WTP is lower than average, yet certain groups are forced by regulations to bear the average costs of public programs to reduce risk, then there will be equity concerns. Indeed, our model permits us to consider the distribution across age and income groups of WTP to reduce the probability of illness profiles of all different kinds.

• A brief comparison of their vsr figures against standard VSL figures would have been useful.

RESPONSE: In the submitted version of the paper, we made this comparison only in the introduction. In subsequent revisions, we made much more detailed comparisons, but later referees requested that we remove all this information. The current version of the main paper contains a moderate amount of information and refers the reader to a variety of published surveys of the VSL literature.

• Judging by the VSI figures for (I yrs sick, then die) and (5 yrs sick, then die) given in Table 4, it appears that in utility terms it is better to kill those suffering from a terminal disease after a period of time than to let them survive. Whilst health states can be judged as worse than death, I'm not aware of any quality of life figures that suggest that even terminal cancer. Certainly revealed preference suggests that most people prefer to go on living.

RESPONSE: It is important to realize that these are ex ante, prospective future health states. Our attitudes about what life might be like should we suffer a future illness can be very different than what life turns out to be like, ex post, if we actually suffer that illness. Yet prospective future health risks are what drive people's WTP for public policies or regulations that will reduce the risk of suffering a probabilistic future illness or accident. Thus we are attempting to characterize people's preferences with respect to future probabilistic health states, not attempting to confirm the quality-of-life perceptions of people who are already sick. It is well known, for example, that paraplegics view their quality of life as better than non-paraplegics predict their quality of life would be should they have the misfortune to find themselves in that state. People adapt to their constraints. We suspect that this referee may be struggling to break away from the intuition invoked in the QALY (quality adjusted life year) literature in health economics, where cost-effectiveness is the main concern, rather than benefit-cost analysis.

7.2 Submission to the *American Economic Review*

We also wish to thank our editor at the *American Economic Review*, where we next sent the paper in 2005. Our paper certainly cannot be described as "short and sweet." It is unavoidably rather complex, given that the data must be understood before the paper can be digested. The first two referees were split. When we submitted our responses to their concerns, they remained split, so the editor went to a new third referee as a tie-breaker. Referee III devoted a huge

amount of time to the paper, identifying what he perceived to be some serious shortcomings. However, his most significant concerns stemmed from misconceptions about things we had not adequately explained in the paper. For example, he thought we had only a single version of our survey instrument, which would have precluded an external scope test. In fact, there were 1801 unique survey instruments, tailored to each individual's age and gender and randomized in terms of attribute levels. We addressed all of his concerns in our revisions, and began to develop this separate document to contain our more-detailed explanations.

However, five months after we resubmitted the paper, this very thorough tie-breaking third referee abdicated from the review process, citing a conflict. The editor then recruited a fourth referee, apparently with a marketing background and experience in the design of conjoint choice experiments. He/she further enlisted a colleague to help with the review. These two reviewers, however, objected primarily to the fact that we had *too many* versions of our survey and had not used a formal blocked experimental design—without realizing that such a blocked design would be inappropriate in our case because the same survey could not be used for subjects of different ages or different genders.

Below, we include the history of review comments by referees I and II in the first round, and by referees I, II and II in the second round, along with our responses in each round, although note that we received no specific feedback about which of our responses the first three referees accepted or rejected in this second round. The editor explained that Referee I had not gotten back to him, he did not contact (positive) Referee II again, and that Referee III had withdrawn.

We also include the combined report of referees IV and V in the third round. They zeroed in on issues that were different from those of the earlier referees, so we conclude that they must have seen our responses to the first three referees. The paper was finally rejected based on the comments of referees IV and V and we were not given the opportunity to set the record straight or to explain why the approaches we had taken in our project were preferable to what was suggested by these referees, or why the results in the paper are minimally sensitive to the alternative assumptions that were proposed. Nevertheless, we include here the replies we would have offered, had we been invited to do so.

7.2.1 First round of reviews from the *AER*, with our replies

The editor's first-round assessment was lukewarm at best, but since the door was left open, we felt obligated to continue. We were influenced by other advice to the effect that the AER is rarely enthusiastic about a paper outright unless they have already decided they will publish it, preferring instead to manage expectations and to provide a justification for a later rejection, should this be necessary, as it is in a large proportion of cases.

7.2.1.1 Editor's first-round comments, with our replies

…One referee asks who the audience for the paper is, and what the key message is (points 1 and 3), and asks also about behavioral approaches (point 2). The referee has a long set of points after that. In a cover letter, the referee repeats points 1 and 3, and says that a revision might be called for but would have a low probability of ultimate acceptance. The referee suggests that a field journal would be a better bet.

Another referee has some questions about substance (e.g., severity of illness) and then a large number of points about presentation and plausibility of the results. In a cover letter, this referee

says that you are to be commended for taking a more comprehensive stated-preference valuation than usual, the "novelty of the approach is exaggerated by claiming that these issues are not well recognized in the field." While I am confident that the presentational issues can be resolved, it is not clear to me that the work is sufficiently innovative to appear in *AER* as opposed to a field journal."

I have read over your paper with these reports alongside. This is not my area per se so I must rely heavily on the referees' views. Their reports are quite different but both have similar bottom lines, which is that there are virtues to the paper but that the probability that a revision would be strong enough for the *AER* is low. I am willing to extend you that opportunity, but urge you to consider carefully whether you should not send your paper to a field journal instead. I would have to send the revision back to these referees, and they are not optimistic (I may send it to a third referee as well, to get another view).

RESPONSE: We have updated the paper to reflect recently published research and we have significantly refocused it to deal more directly with the three major concerns raised by both reviewers.

1. **Novelty of our contribution.** We make it clearer that we propose a new approach to measuring the benefits of prospective mortality risk reductions—one that we argue should replace the Value of Statistical Life (VSL) approach which has dominated the literature for twenty years. This new approach places new demands on researchers but also yields more theoretically sound, more general, and more transparent benefit measures, especially for the value of marginal sick years and lost life-years.

2. **Relevance for a general audience.** We also make it clearer how this research will benefit not only researchers in the fields of health, environment and safety, but likewise those in applied labor and public economics who focus on major lifecycle decisions that depend upon people's expectations about their future health states.

3. **Validity of our results.** We present clear evidence that our results are comparable to those generated in the narrowly defined special case addressed most commonly in the existing literature. Moreover, for those who remain concerned that poorly executed stated-preference results will exaggerate benefits estimates, we point out that our estimates actually tend to fall somewhat below most conventional reveal-preference estimates for the analogous constructs.

We have also added some new simulations which will be of great interest to the constituency which worries about VSLs as a function of age. Most researchers find that estimated VSLs follow an inverted-U shape as a function of age, peaking in mid-life. We demonstrate the same relationship for our special case of "sudden death now." But our model also readily allows age profiles for estimates for WTP to reduce the risk of other illness profiles. We were already illustrating the case of "ten years of latency, five years sick, followed by death." Now we also include estimates of WTP to reduce the risk of "six months of illness near the end of life, followed by death six months earlier than otherwise." The age profiles for this

WTP decline slightly from age 25 to age 60, but then begin to rise sharply. No other modeling framework currently provides WTP estimates for these types of scenarios.

7.2.1.2 First-round comments of Referee #1, with our replies

General Comments:

(1) The intended audience for this paper is not clear.

REPLY: There are two major audiences for this paper. First, in the fields of health, public and environmental economics, this information is essential to value accurately the benefits from medical research (Murphy and Topel, 2006), or from environmental, health and safety regulations (Viscusi, 1993). Second, labor economists are also interested how these policies or programs improve individuals' expected future health states because these expectations affect widely studied life-cycle decisions related to consumption and savings, as well as participation in labor and health insurance markets. We make this clearer in para.1, pg.1 of the introduction.

Over the last ten years the audience at the AER and other leading general-interest journals seems to have been interested in many aspects of this topic. Included among the key references in our paper are the following earlier works:

Kniesner, T. J. and W. K. Viscusi. 2005. "Value of a statistical life: Relative position vs. relative age." *American Economic Review* 95(2): 142-146.

Ashenfelter, O. and M. Greenstone. 2004. "Estimating the value of a statistical life: The importance of omitted variables and publication bias." *American Economic Review* **94**(2): 454-460.

Cutler, D. M. and E. Richardson. 1998. "The value of health: 1970-1990." *American Economic Review* 88(2): 97-100.

Dow, W.H., Philipson, T.J., and Sala-I-Martin X. 1999. "Longevity complementarities under competing risks." *American Economic Review* 89 (5): 1358-1371.

Ashenfelter, O. and M. Greenstone. 2004. "Using mandated speed limits to measure the value of a statistical life." *Journal of Political Economy* 112(1): S226-S267.

Murphy, K. M. and R. H. Topel. 2006. "The value of health and longevity." *Journal of Political Economy* 114(5): 871-904.

Hall, R. E. and C. I. Jones. 2007. "The value of life and the rise in health spending." *Quarterly Journal of Economics* 122(1): 39-72.

Smith, V. K., M. F. Evans, H. Kim and D. H. Taylor. 2004. "Do the near-elderly value mortality risks differently?" *Review of Economics and Statistics* 86(1): 423-429.

(2) Is the main purpose to sell an approach that has been used in other fields to economists? If the latter, why should economists outside such fields as environmental, health, and perhaps labor economics be interested? And even such readers will have questions about specific details.

REPLY: Our approach is original to us. It has not been used in other fields. The main purpose of the paper is to present a new approach to conceptualizing, and then empirical measuring, preferences over inter-temporal health risk reductions. The literature has long complained about several simplifications and limitations associated with the VSL. Our approach provides a new but still rigorously utility-theoretic alternative to the VSL.

While researchers have long recognized the limitations of the VSL approach, they have been unable to overcome them because of the constraints of existing empirical data and methods. The improvements identified in this paper take the form of 1) enhanced construct validity (i.e. that the measure varies systematically, in ways we would expect, with attributes of the health threat and characteristics of the individual), 2) utilization of illness profile information that is omitted from the conventional estimation framework, 3) enhanced transparency of the determinants of demand through an explicitly structural model, and 4) replacement of the concept of a one-size-fits-all VSL with a measure of heterogeneous individual benefits: willingness to pay (WTP) for a risk reduction for a specified health threat.

There would be little in this paper that was noteworthy if our method offered merely one more estimate of "the VSL." What makes this paper important is that "the VSL" is only one very special case of the WTP functions that our data and our model permit us to estimate. In one seamless model, we provide policy-makers with an opportunity to escape from the following troubling (although politically popular) limitations of many earlier models:

a.) a single *one-size-fits-all* value of a statistical life, focused only on mortality risks (e.g. "sudden death in the current period") regardless of the type of risk or the characteristics of the affected population;
b.) a single VSL that makes no adjustments for the *quantity* of life lost, so that willingness to pay to reduce risks to young people is constrained to be the same as willingness to pay to reduce risks to the elderly (where previous ad hoc adjustments have tended to involve dividing a single VSL estimate by the average number of discounted life-years remaining to produce the "value of a statistical life-year.");
c.) insufficient attention to fatal versus non-fatal illnesses, and *pre-mortality morbidity* (separate models have been used to attempt to identify the value of a statistical illness, but these illnesses have been specific named illnesses, rather than a continuum of possible illness profiles);
d.) VSL estimates which are independent of the income levels of the affected populations. These measures are *inverse demand functions*, no more and no less. As such, one would expect that income would have a systematic effect on willingness to pay.

It is difficult to argue that there is any topic more relevant to the general population than "life and death." We propose and demonstrate a major improvement to the methods by which we attempt, as economists, to assess the benefits of a wide range of public and private programs to reduce risks to life and health. Every one of us, economist or not, is impacted by these methods, since these assessments affect the stringency of so many regulations with which industries must comply and thus the prices of so many products

we buy. Perusing any past issue of the AER reveals many topics which are much more narrow and specialized. We have made an effort, in our revision, to spell out the "sufficiently general interest" case much more plainly in the introduction.

(3) Some economists are quite critical of the type of approach used in this study. Specifically, how do the results show that these criticisms are invalid, if they are involved?

REPLY: Stated preference studies can indeed be conceived and executed badly. Many economists still base their impressions of stated preference research on what they once read in a symposium in the 1994 issue of the *Journal of Economic Perspectives* devoted to the opinions of several participants on both sides of the Exxon Valdez litigation in the early 1990s. In contrast, the survey upon which our research is based uses state-of-the-art strategies in survey design, thorough pretesting, the standing representative consumer panel of one of the premier survey research firms, careful assessments of construct validity and scenario rejection, and rigorous econometric analyses. Anticipating the usual lingering concerns from the Exxon Valdez era, we went to great lengths during pages 7-13 of the original paper to explain the degree of care that was taken in these steps. A particular concern of ours is that of hypothetical bias, as has been documented by Cummings and Taylor (1999) and List (2001) and a remedy shown to exist by List (2001). As a correction, we employ a similar remedial strategy in the forms of a "cheap talk" reminder—to ensure that respondents carefully consider their budget constraint and to discourage them from overstating their willingness to pay.

Given that revealed preference data to address our questions is not available—and never will be—these data represent the best source of information to date for the tasks at hand. Attention to stated preference data for VSL estimates is also not new. For example, the U.S. EPA currently relies on a meta-analysis including both revealed and stated preference estimates to guide its choice of a VSL to use for policy analysis and regulation.

(4) The author(s) packs a lot of detail into this paper. Yet results are not explained. Are specific findings plausible or not? There may be too much here for one paper. Reducing some repetition would only cut the paper about a page.

REPLY: We are now careful to explain our central results. See pgs. 33-38. Myriad different empirical estimates of the value of a statistical life exist in the literature (so many, in fact, that a number of meta-analyses have already been conducted). Thus we can carefully cross-validate and assess the plausibility of our findings. For the special case of "sudden death in the current period," our findings can be compared to conventional estimates of the "value of a *statistical* life" (VSL). The VSL was contemporaneously estimated by the U.S. EPA to be about $6.2 million, although smaller numbers have been used by other federal agencies. By comparison, for a 45-year-old with $42,000 in income, using a 5% individual discount rate, our model produces an analog to the VSL with a point estimate of $5.35 million and a range from $3.56 million to $7.43 million. This would seem to answer the "plausibility" question. The paper now goes into much more detail in comparing the implications of our model with the quantitative findings of other

research in the directly comparable case of “sudden death in the current period,” both overall and as a function of the respondent’s current age. But the important contribution of our research is that we can also derive willingness-to-pay estimates for a wide array of *other* illness profiles.

Specific Comments:

p.3. Explain how effectiveness parameters included.

REPLY: We are not sure what you mean be effectiveness parameters. We did assess individual subjective assessment of the effectiveness of the interventions. However, we decided not to complicate our analysis (and lengthen the paper any further) by exploring this source of heterogeneity in this paper.

p.3. There are many criticisms of this type of research. In economics, the critique now falls under the heading "behavioral economics." The following types of issues are reused [sic].

(1) Can people deal with such complex choices/bounded rationality"?

REPLY: People are routinely faced with, and must often make, precisely these types of complex decisions—especially as they age. So the operative question is not so much “can” people deal this these choices but “how well” do they do so. Individual decision making may be influenced by biases and heuristics. Nonetheless, in making these decisions, people are revealing what they perceive to be their expected benefits. We have taken great pains to recover these preferences while minimizing or controlling the effects of complexity, biases and heuristics. For example, we have sought to minimize informational loads while still retaining a plausible and contextually accurate choice environment. We test for order effects. We have the ability to control for anchoring on prior information on the respondents possessed about each type of disease risk.

(2) People may be myopic and not forward-looking.

REPLY: We agree that this may be true of some people. But we offer two responses. First, our empirical analysis reveals that this is not true of the average person in our sample. The statistical significance of the estimated coefficients on the health-state duration variables reveals that respondents are indeed attentive to changes in their future timeline of health risks. Second, even if some people are atypically myopic, their demand for these programs is still relevant in a positive analysis which measures the aggregate ex ante social benefits of these types of programs (i.e. for health risk reductions at some cost).

(3) The assumption of exponential discounting may be violated. These arguments are briefly addressed later in the paper, but the discussion is too brief.

REPLY: While the assumption of exponential discounting may be violated, it should be noted that no prior research has comprehensively entertained discounted expected

durations in sequences of different future health states as the explicit object of choice in a study of this type. It is certainly possible to consider hyperbolic discounting, and indeed we do consider this alternative in other similar work. However, considerably more infrastructure would need to be discussed, as well as many more parameter estimates, if we were to pursue comparisons of alternative models of individual-specific discount rates in this paper. It would require the discussion of a separate survey of another sample from the same population, the specification and estimation of an individual-specific discounting model, and the transfer of that model to this sample based on the different measurable characteristics of each individual. It would add a very large amount of detail to the story, which is why we plan to cover that angle in a subsequent paper, after the basic model has been laid out here. In the present paper, we demonstrate sensitivity to different fixed assumptions about an exponential discount rate. As this reviewer states above, there may already "be too much here for one paper."

p.4. There is a lot of heterogeneity in health risks, both because of variability in genetic makeup and health behaviors. Would not a lifelong smokers or an obese individual value these programs differently from a never smoker or a person of normal weight? How is much variability reflected in the study?

REPLY: Yes, objective and subjective health risks present an important source of heterogeneity in the demand for health-risk reductions. In a separate paper, based upon the working model derived and illustrated in this paper, we allow the utility parameters to vary systematically with the individual's current *objective* health status (prior or current experience with the same illness that serves as a label for each offered program, as well as the number of other major illnesses from which the individual suffers). We also allow these utility parameters to vary systematically with the individual's *subjective* risk rating for the same illness and his/her average rating for other illnesses. In yet another paper, we differentiate our willingness-to-pay estimate by illness label and interact these with other health habits such as smoking. To extend the analysis in this paper to include this heterogeneity would lengthen the paper considerably, which is why we elected to limit the scope of the current paper to the development and explanation of the basic model. This basic model features some essential heterogeneity with respect to age (mimicking this key source of variation in other previous work). Again, as this reviewer states above, there may already "be too much here for one paper."

We now note specifically in the paper that suppression of heterogeneity across types of illnesses and across types of individuals (beyond their ages and income levels) means that the utility parameters reported here can be considered the average levels across the menu of these major illnesses and injuries, with the range of symptoms and treatments used to describe each one. The randomized design of the stylized illness scenarios (subject to minimal plausibility) means that illness labels are essentially independent from the descriptions of illness profiles as well as the characteristics of the respondent, so there is minimum omitted variables bias from suppressing this type of heterogeneity.

p.4. Explain Graham's option price. This is mentioned several times in the paper. Discuss once and describe well.

REPLY: We now define the Graham (1981) option price precisely at the beginning of the modeling section.

p.5. Explain latency period. What, for example, is the latency period for Alzheimer's disease? Latency is used in the tables, but never explained. Nor is it clear where the values come from. Is latency really known for these diseases?

REPLY: The concept of latency is now defined when it is introduced. Latency is defined as the elapsed time between now and the onset of disease symptoms or the occurrence of the specified accident. Latency is a crucial aspect of most environmental health risks. Few of these risks kill instantly. Pre-testing revealed that respondents were generally aware that they face many types of health risks that will not fully come to bear for many years, and that the exact timing of the beginning of symptoms was implicitly an expected value for a distribution. However, we specified the start date of the illness (or the date of the injury) as a particular number of months or years into the future, since any more detail seemed to create information overload. We have tried to make it clear in the paper that the expected illness profiles for each choice set are not tied to any actuarial patterns, since most illnesses are not reportable. Cause of death and age at death must be reported as vital statistics, but not age at the onset of illness or the pattern of illness. We randomized the illness profiles from a range of plausible possibilities, and carefully culled those which nobody would believe.

p.5. bottom. A lot of people (perhaps half or more) do not receive recommended screening, at least at recommended rates. Is this reflected in a high frequency of low WTP for screening in this study?

REPLY: It depends on the reasons why people do not receive the recommended screening. If people do not view the risk information and the resulting therapies as effective then they might not be willing to pay for them. During the pretesting phase of this study, we dropped those illnesses for which people did not think screening could yield valuable information in general.

Second, respondents might view screening as effective but not feel they are truly at risk from the targeted disease—or they may view themselves as at risk but feel they face greater risks from other illnesses. This is a legitimate economic reason for being unwilling to pay for such a program. Subjective risks from the same and other illnesses, as determinants of willingness to pay for specific risk-reducing programs, are explored in a separate paper.

Third, for afflictions which are considered to be readily avoidable, like skin cancer for example, people are much less willing to pay for screening programs. In these cases, the modal willingness to pay is even zero in some cases. Our models do not constrain WTP to be strictly non-negative, and respondents are not given the option to express negative WTP, so portions of the fitted probability density for WTP which lie in the negative domain should be interpreted as zero.

p.6. on sample representativeness. Compared to which data? Discuss only once in this paper, not twice and incomplete both times.

REPLY: We agree and now discuss this only once. We are also careful to discuss the comparable revealed preference (RP) data more specifically, since wage-risk studies typically apply only to working-aged adults, and often only to males. Our sample, drawn as it is from the standing consumer panel of Knowledge Networks, can be expected to be much more representative of the U.S. population in terms of observables. Respondents to our survey range in age from 25 to 93 years old, and we include both males and females, in the labor force or not.

A greater concern with survey data, however, is sometimes the representativeness of the sample in terms of unobservables. Given that our "outcome" equation is a three-alternative conditional logit model, conventional two-stage Heckman selectivity-correction models are inappropriate. One could estimate a selection equation and construct an inverse Mill's ratio to use as a regressor in the second stage, but conditional logit errors are, by definition, uncorrelated with anything, so the intuition of the conventional Heckman correction model does not carry over. Instead of this strategy, we instead fit a selection-model to predict presence in our estimating sample for each of the over 525,000 initial random-digit-dialed (RDD) recruiting attempts. Nobody, including any other user of Knowledge Networks' consumer panel, has attempted this type of comprehensive response propensity correction. We are the only researchers with whom Knowledge Networks has shared its data about initial RDD telephone contacts.

We control for heterogeneity in comprehensive response propensities from this RDD pool by allowing each estimated utility parameter to vary systematically with the *deviation* of each respondent's response propensity from the average in the pool of 525,000 initial RDD contacts. Only one parameter exhibits significant heterogeneity in this dimension. We use its predicted value for an individual with the mean response propensity. Although one referee suggests we just drop this correction since the effects are relatively small, we would prefer to retain the correction, since its effect is statistically significant and the model is clearly mis-specified without it.

p.6. Although you undertake validity tests, you do not assess whether specific values are plausible. For example, is the estimated value of $2.35 million to avoid one year of nonfatal illness at all plausible (see Table 4)? To this reviewer, it seems implausibly high. If this is the value, we will exhaust GDP in preventing such outcomes.

REPLY: We offer a couple of points in response.

This reviewer's comment highlights the public relations fiascos that seem to occur with alarming frequency whenever the subject of valuing mortality risk reductions appears in the press. The use of the standard "statistical lives" terminology in the previous version of the paper was not our choice, but the accepted convention in the literature. At some point, someone decided to average different estimates of the WTP for risk reductions

(actually a marginal rate of substitution between money and risk reductions) by first normalizing on an arbitrary risk reduction of 1.00 (whereas most empirical risk reductions actually under consideration were in the range of 10^{-6} to 10^{-5}). The resulting "value of a *statistical* life" (VSL) is not the same thing as "willingness to pay to save one life with certainty," or "the value of a human being," although the general public is often inclined to think that these things are the same.

This confusion may actually have contributed to Governor Christie Whitman stepping down as the EPA Administrator shortly after the EPA attempted to differentiate the benefits of health risk reductions for different age groups by using a VSL for seniors that was 2/3 of that for adults. Another round of moral outrage flared after July 10, 2008, when an Associated Press reporter broke a story about the U.S. EPA devaluing human lives. Non-economists do not appreciate the important difference between an inverse demand measure in the form of "aggregate willingness to pay for small reductions in mortality risk" and the philosophical question of the "worth of a human being." (See the backlash inventoried by Trudy Cameron at http://www.uoregon.edu/~cameron/WTPM/VSL_confusion.pdf)

We have therefore decided to take a bold step in our revisions to this paper. We will attempt to nudge the literature away from the concept of "the VSL" and steer it towards using just the WTP for individual risk reductions. Following Howard (1984), we will preserve the traditional assumption that WTP is roughly proportional to the size of the risk reduction, but change the normalization from a 1.00 risk change to a different risk change that is more in the realm of many policy questions—a 0.000001 risk change, dubbed a "microrisk" by Howard. This implies no change in our estimates. It means only that instead of reporting the "value of a statistical illness profile (VSIP)" in millions of dollars as we were doing in our original draft, we will report the "WTP for a microrisk reduction for the specified illness profile."

With respect to our results for morbidity scenarios, nobody is being asked to pay $2.77 million to avoid just *one year of nonfatal illness with certainty*. Instead, people's choices imply that they would be willing to pay about $2.77 to avoid a 0.000001 *chance* (i.e. a microrisk) of this whole illness profile (or $277 to avoid a 0.0001 *chance* of this illness profile).

It is also important to realize that these are major illnesses, and that respondents clearly do not view "recovery/remission" as a return to their prior status-quo health state. Otherwise, the coefficient on the *pdvr* term would be zero. These are major illnesses, including several types of cancers and most of the other things from which we most often die. It was not plausible to tell people that there would be 100% recovery should they experience one of these major illnesses, so we did not urge them to assume this. The implied value of a marginal discounted sick-year is inferred in a hedonic fashion from respondents' choices across a wide variety of illness profiles. We believe our results allow us to value particular illness profiles within the range of profiles we offered, but we would hesitate to apply our results out-of-context—for example, to a single sick-year in isolation with 100% recovery.

p.6. The third point seems like an oversell. There are very many substitute risks and the relevant ones vary among individuals. For example, a never smoker will not worry much about getting lung cancer.

> REPLY: The broader point we seek to make here is that we actively encourage individual to reflect upon a much broader set of risks among which they may substitute expenditures. Prior studies in the literature often focus upon only one, or at most two, sources of health risks. We have altered the language of this section to try to better reflect this.
>
> Incidentally, in a separate paper where we differentiate preference parameters by illness labels, we do indeed find that non-smokers are far less concerned about lung cancer and respiratory disease. Thus this reviewer's intuition is borne out.

p.6, bottom. Limit discussion of the survey to items and in this analysis. For example, are subjective risk assessments used in this paper's analysis?

> REPLY: We agree with your central point and have limited detail discussion to only the analysis presented in the paper.
>
> However, subjective risk assessments, mentioned in our paper's review of the structure of the survey, serve an important role even if these variables are not used in the analysis. The elicitation of subjective risks is intended to induce the respondent to begin thinking about the wide range of threats to life and health and the different sorts of illnesses that may afflict them during their lifetimes. They will not be considering just one threat, but ten different major illnesses or injuries, and we do not wish to exaggerate the salience of the first illness profiles they will consider by failing to get them thinking, in advance, about the existence and relevance of many different health threats.
>
> We are also eager not to leave readers with the impression that the basic model which we demonstrate in this paper is the most that can be possibly be done with the available data. Subjective risk assessments are useful to demonstrate the construct validity of our findings, and their influence is addressed in other papers which explore the many different questions that can be addressed with our data. We could go down that path here as well, but again, we agree with this reviewer's earlier opinion that there may already "be too much here for one paper."

p.9, top. Programs to manage. Again this may be a bit of an oversell. What for example can be done to manage Alzheimer's disease? Or are these hypothetical programs used in the survey to elicit WTP

> REPLY: Yes, these are hypothetical detection and prevention programs. These programs are now specifically identified as "early diagnostic programs." Pages 17 through 24 of the survey instrument explain these hypothetical programs and exactly how they should be viewed by the respondent. These are ex ante risk reduction programs

pitched as “pin-prick blood tests once a year. Each test works by checking for chemicals in your blood to indicate you are at risk for an illness. 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.” (Of course, programs to reduce the risk of traffic accidents are described differently.) These minimally invasive early diagnostic programs were used because they seemed the most plausible and acceptable to the majority of test subjects during our several cycles of survey development.

New treatments to detect and manage the onset of even Alzheimer's disease now exist. See http://www.webmd.com/alzheimers/guide/treatment-overview

p.9. and Table 1. The reader needs help with Table 1 (and other tables). At present, tables are largely just included, forcing the reader to draw his or her own inferences. Where do the values in Table 1 come from? How did you come up with 4.0 illness years for a traffic accident for example? What is a latency period of 18.2 years for traffic accident? For some accidents, the latency period may be 15 minutes or less. Same point for recovery period, chronic effects, etc.

REPLY: Table 1 is now more specifically labeled as the “Range of Attributes used for Stylized Illness Profiles.” The illness profiles are not the REAL distributions for these diseases, but the outcome of a randomized design process. Orthogonal mixes of illness attributes were generated. Then a randomly assigned label was associated with the type of illness/injury. Implausible combinations were then eliminated. Our goal was to span the range of possible illness profiles in each illness/injury category, but not to replicate the actual distributions—this information proved impossible to find, anyway. We require only that respondents treat the stated latencies, sick-years, recovered/remission years, and lost life-years as the “expected values” of the joint distribution of these states, should the illness/injury in question strike them.

Fortunately, early interactions with test subjects revealed that ordinary citizens seem to have very little understanding of the real actuarial profiles of particular illnesses (unless they have personal experience with the illness, either themselves, or in a close family member or friend). We elicit these experiences and intend to use them as controls in yet another subsequent paper.

The pool of illness profiles used in the elicitation of preferences is therefore unrepresentative of the true distribution of illness profiles in the population. Since we depart from the true joint distributions of health states which make up real illness profiles, it subsequently makes little sense to simulate sample mean WTP for the configurations of illness profiles used in our stated preference scenarios. Instead, we illustrate with selected benchmark cases.

p.12. First paragraph needs editing. Say more about cognitive complexity.

REPLY: With respect to cognitive complexity, we recognized the importance of this at the very beginning the of survey instrument design process. Concerns about

"information overload" guided our selection, definition and presentation of attributes. Inevitably, given the finite duration of the survey, some tradeoffs had to be made.

We continued to evaluate choice difficulty even during the survey itself by including a difficulty rating after each choice scenario in the survey. Responses to this question have been analyzed in yet another paper.

Perhaps the greatest indication that we effectively address this concern is that individuals clearly recognized and reacted to differences in all of the core attributes analyzed in this paper; the coefficients on every one of them are statistically significant.

p. 12. Say more about the budget constraint. Future health shocks could have major effect on the household's budget constraint, and some people have preexisting illness.

REPLY: We agree. In a separate paper we explore the effects on WTP of both subjective expectations about the probability of future illnesses as well as pre-existing illnesses.

Our empirical model permits alternative assumptions about the proportions of income that people assume they will earn while in each health state, but the data provide no basis for choosing any one proportion over any other, beyond what we have already implemented. If another assumption seems more supportable, we can certainly do a sensitivity analysis. It proved too difficult to attempt to elicit each respondent's subjective future income profiles, conditional on each probabilistic illness.

p. 13. Say more about the conjoint analysis findings. Just referring to Table 2 is not enough.

REPLY: The paper now states the following: "In Table 2, Model 1 demonstrates that even minimal conditional logit choice models, in terms of the raw program attributes, produce intuitively plausible and strongly significant coefficients on the two most crucial aspects of each program: a lower cost and a higher risk reduction make a program more attractive. Model 2 shows that the other two most important dimensions of the illness profiles, the number of sick-years and the number of lost life-years for which the risk will be reduced, are also strongly significant determinants of respondents' choices among programs. Respondents are systematically more likely to choose programs which address more serious health threats."

Table 2 fulfills the usual initial "face validity" requirement—that stated preference analyses should demonstrate "sensitivity to scope" in some very basic specifications before the researcher embarks upon more complex models. Models 1 and 2 demonstrate that preferences among programs and the status quo are significantly affected by the cost of the program and by the durations of the adverse health states in the illness profiles for which the programs are intended to reduce risks. This is shown in the pair of ad hoc specifications (Models 1 and 2). It is subsequently shown for a model that is based on an underlying utility-theoretic structural model (Model 3, discussed after the structural model has been introduced) that depends upon the square root of net income and is additive and linearly separable in the amount of time in each adverse health state. With

this requisite sensitivity to scope duly demonstrated, we move on, in Table 3, to the more-general models that appear to be warranted by the data.

p. 13. Discuss the sample when this is mentioned the first time.

REPLY: The survey sample is still mentioned first in the introduction, where we spend three paragraphs introducing it. The sample is then described in considerable detail in the five pages of Section 2 (Survey Methods and Data), commencing directly after the Introduction and preceding the development of the utility-theoretic choice model in Section 3.

p. 14. Many illnesses do not have a post recovery state. The condition is chronic for life, for example diabetes, Alzheimer's. How was this handled?

REPLY: After the randomization of illness profiles, we excluded implausible cases, such as recovery from diabetes or Alzheimer's disease. As shown in Table 1, there are no instances of sudden death or recovery from diabetes or Alzheimer's disease (rows 7 and 8 of the table, last two columns). Sudden death occurs only for heart attacks, stroke, and traffic accidents.

p.14. Illnesses affect the budget constraint, due, e.g., to work-loss. How was this handled?

REPLY: We model individuals as assuming that disability insurance will make up their income to current levels. It proved too daunting to attempt to forecast individual subjective income profiles as well as health profiles. We exceeded our permitted survey duration in far too many cases as it was.

p.17. Life tables differ among individuals within age and gender categories. Values can be assumed in the survey, but respondents may find these values implausible. How did you deal with this issue?

REPLY: Many of our test subjects found their actuarial life expectancy to be implausibly short. Instead, many people seem to talk about the average age at death of parents or grandparents who died of "old age," as opposed to dying prematurely from heart attack, cancer, or other acute illness. We found that adding eight years to the actuarial life expectancy for someone that age produced a prediction that appeared plausible to most individuals. We added one year for females and subtracted one year for males.

p.20. "Respondents seem not to interpret being recovered.. ." Judging from which results?

REPLY: This is based upon the finding that the estimated marginal utility from a discounted recovered/remission-year is not the same as the (zero-normalized) utility from a pre-illness year.

Consider an illness profile with no lost life-years, so that all of the $\log(pdvl_i^j + 1)$ terms in Table 3, Model 7, are zero. Assume also that the individual's propensity to participate in

the survey is typical for the Knowledge Networks RDD recruitment pool. The fitted marginal (dis)utility associated with the shifted log of discounted sick-years is -50.52. The fitted marginal (dis)utility associated with the shifted log of discounted recovered/remission-years is only -17.09. This implies less *dis*utility from recovered/remission status than from sick-time, but not the 0 marginal (dis)utility that would correspond to restoration of the pre-illness health state.

As mentioned above, respondents were left to assume what they wished about their health status in the post-illness period, if the illness profile specified that this disease would be non-fatal. For illnesses as serious as our labels imply, it was not credible to assert that recovery/remission would mean "complete recovery to the individual's current status quo health state." We thus infer the utility of time in the post-illness recovered/remission state, relative to current pre-illness utility, based on people's average implicit assumptions about their likely post-illness health state. An important implication is that we cannot really simulate WTP for an illness profile that involves a non-fatal illness with "recovery/remission" without carrying along the post-illness utility decrement for each year in the post-illness state.

p.21. It is not clear a priori that marginal utility should decrease in health state years, at least it should not be monotonically decreasing.

REPLY: Models which are linear in discounted health-state-years fit less well than models which are log-linear (actually, partially translog) in discounted health-state-years. The behavior of the marginal utility of a discounted health-state-year is an empirical question, and these data appear to suggest (strongly) that marginal utility of a discounted adverse health-state year is decreasing, although when we generalize to model with interaction terms—such as those featured in this paper—the number of discounted sick-years preceding death has a statistically significant effect on the marginal utility of a discounted lost life-year. If you are sick long enough, the prospect of being dead is less unpleasant.

It is important to bear in mind that we do not attach a lower utility to each additional chronological year of sick-time in a given spell of illness. The model merely says that when individuals contemplate scenarios with a longer spell of illness, versus a shorter spell of illness, the overall disutility of the longer spell is greater, but not by as much as the average disutility of a year of time in the shorter spell. We do not ask individuals to value successive years of illness in a given spell, one at a time. We only ask them to make tradeoffs with respect to spells of different stated sizes.

p.22. The improvements in log-likelihood seem small, not substantial.

REPLY: Relative to what? Changing to the logarithmic form improves the log likelihood from -11726 to -11720, retains the strong significance of all parameters. This change also allows the sick-years, recovered/remission-years, and lost life-years coefficients to take on the relative values of -27, -23, and -29 (which may seem to be a somewhat more plausible ranking). Subsequently adding the quadratic term in lost life-

years and the sick-years/lost life-years interaction, as well as the heterogeneity with respect to the respondent's current age, improved the log likelihood to -11685.

p.24. Cite literature on time preference, include time preference in the health domain. Does the literature suggest discount rates that would be appropriate for this analysis?

REPLY: In a footnote, we now cite the Alberini et al. (2006) results, for Canada and the U.S., for the implicit discount rates associated with willingness to pay to avoid latent health risks. There is, as yet, no consensus on what is the "typical" discount rate which should be applied to future health states as opposed to future money.

p.28. Discuss Table 4. Do the estimates other than for VSL seem plausible? It is not at all surprising that the assumed discount rate affects the results.

REPLY: We now spend a paragraph discussing the main results in Table 5 (the middle column), with the results in the other columns serving as sensitivity analysis with respect to the assumption about discount rates.

7.2.1.3 First-round comments of Referee #2 and our replies

This paper presents a conceptual model and empirical analysis for estimating individual's willingness to pay (WTP) to reduce risks of fatal and non-fatal illness. Although the fact that WTP might depend on characteristics of the illness, life years lost to death, and other factors has been well recognized and discussed for decades (see, e.g., Jones-Lee, Hammerton & Philips 1985 for an early attempt at empirical estimation), the present paper provides an integrated treatment within a single model that is arguable better than what has been offered before.

REPLY: We have added the Jones-Lee et al. (1985) paper to the references. We have also made it clearer that we not simply identifying new limitations of the VSL approach but rather offering a simpler and more general approach that addresses many long-recognized limitations of the conventional approach.

The conceptual model assumes that individuals choose between binary lotteries on health profiles in accordance with expected utility, where the utility of a health profile is assumed to be a function of the discounted number of years of life lost, years of illness, years before illness, and years after recovery from illness (or years in remission). The marginal utilities associated with each class of years are allowed to vary with the individual's age at the time he chooses between lotteries. Compared with some of the other literature in this area, a key limitation of this approach, at least as implemented in this paper, is that it does not distinguish between the severity of illness, i.e., all years when ill have exactly the same marginal utility (except for the effects of discounting and the age-related shifters just mentioned). (1) Since the severity of illness can vary substantially, this is clearly a significant limitation. Moreover, the authors do not provide any information about the severity of the illnesses presented to survey respondents other than giving us the names of the illnesses and one example (Appendix A). (2) In addition, the authors impose the constraint that future years of life income are discounted at a common rate, which seems unlikely, but is probably less important than (3) the maintained assumption about adding years in each state.

REPLY: We address your three concerns in turn.

(1) We agree that the severity of illness should be expected to be an extremely important attribute of the individual's health state. In fact, our choice sets indicate the pain and disability associated with the major illness or injury featured in each different illness profile, characterizing it as either moderate or severe. We also describe whether minor or major surgery would be needed and how long the periods of moderate and/or severe pain and disability would persist. During the tutorial portion of the survey, which works carefully through each distinct feature of the upcoming initial choice set, page 14 of the survey goes into considerable detail about how respondents should interpret these moderate and/or severe designations for the pain and disability associated with an illness or injury. Specifically:

Each illness may cause pain and disability. Below we describe what it is like to experience moderate and severe pain and disability.

Color key for level of PAIN and DISABILITY:

Moderate	Pain: Some discomfort performing daily activities; most pain can be controlled by medication. Disability: Some problems walking, washing, dressing or using the toilet.
Severe	Pain: So bad it impairs daily activities. Difficult to control even with medication. Disability: unable to perform usual daily activities; usually confined to bed; unable to wash, dress, or use toilet independently; unable to communicate well with others.

The characteristics of the stylized illness profiles are randomized to the extent possible, subject to plausibility, in order to minimize the threat of omitted variables bias if these attributes are omitted from a specification.

In this paper, we chose not to explore the effects of pain and disability levels, since this additional dimension would add yet another layer of complexity to the analysis. Instead, we decided to focus on the most basic features of each illness profile in this initial paper, and to pursue generalizations that emphasize additional sources of heterogeneity in subsequent papers which will refer to this basic analysis. We rely upon the randomization of the omitted illness profile variables to minimize any potential bias. Of course, such bias would likely be a greater concern had there existed any data on real disease profiles that we might have used. Unfortunately for analysts, most illnesses are not reportable.

(2) Our model requires that we assign a discount rate which could in principle be individual-specific. However, in designing the survey, it proved difficult to ascertain from individual respondents what differing discount rates they might employ over different time horizons, and to future health as opposed to future net income. We would be happy to entertain any suggestions you may have as to what rates we should employ.

In separate survey of a different sample from the same consumer panel, we elicited information about each person's financial discounting preferences. We are in the process of attempting to transfer the discount rate functions from that sample to the estimating sample for this paper in the form of fitted individual discount rates. However, the unexplained variation in individual-specific discount rates seems to remain high. Preliminary results suggest that this much more complicated model does not perform appreciably better than the one used in this paper.

(3) Our health profiles do not consist of "adding years in each state." Instead, they are cast as changing the *probability* that the individual might experience an adverse health profile consisting of specified intervals in each of three adverse health states, relative to the health profile they might otherwise experience. (We wish we could be more responsive to your underlying concern here, but we are not exactly sure we understand what it is.)

I have a few concerns about this paper, related to clarity of presentation and plausibility of the empirical results.

Presentation

First, the theoretical model and empirical equations are rather difficult to follow, in part because a lot of new notation is introduced. Unless I missed it, some of the terms are not defined. One example is what I take to be changes in probability **delta pi.** In interpreting the empirical results, the sign of this variable is critical since coefficients on a variable equal to the product of this change in probability and discounted numbers of years in various health states are estimated. Also, the variable "risk reduction" used in Table 2 is not defined. Why is the subscript S,H required on the expectation operator in equation 2?

REPLY: The risk reduction "delta pi" term is now defined immediately after we introduce the constituent without-program and with-program risks. The risk reduction variable in Table 2 is made explicit as the absolute size of the risk change, so that in naïve models a program that offers a larger absolute risk reduction is more attractive to respondents, ceteris paribus. The subscript S,H has been dropped from the expectation operator and we have merely mentioned in the text that expectations are taken across the sick and healthy states. Also, while it makes no substantive difference due to the constancy of all variables within any given health state, we have made it explicit that the quantities under consideration in the choice model are present discounted values of expected utilities. We take expectations first, then discount.

I believe that variables like *pdvi* represent the present value of the years when ill (conditional on becoming ill given the health profile specified for a particular intervention). However, it is not clear to me what the counterfactual is. If the respondent does not suffer the specified illness, is he assumed to remain perfectly healthy until dying in ti years (where ti is his "nominal life expectancy")? Alternatively, the respondent might imagine a portfolio of health risks he may face as he ages, and incorporate this portfolio into both arms of the binary lottery described by "program A" (i.e., with or without illness). This issue is important for interpreting the estimated marginal utility of a year in ill health - is it compared with a healthy year, or compared with a year with some expected health condition that would reasonably depend on the age at which it is experienced? It is also relevant to footnote 32 on p. 20 which says that a positive value for the estimated marginal utility of a lost life year would imply that illness is worse than death - I would have thought such a value implies that a healthy life year is worse than death.

> REPLY: The individual is not choosing between sickness and health. We now make this explicit in the paper, in the paragraph following equation (1). They choose between the offered programs and the status quo. We have added a passage that reads: "A key insight is that individuals are informed that they have an existing risk of suffering from the illness or injury in question. Their choice is not between suffering from the illness and enjoying perfect health, since there is a specified chance of suffering the illness both with and without the program. Instead, their choice concerns whether to purchase a program that will reduce their risk of suffering from the illness in question by a specified amount."
>
> The respondent faces a current status quo probability of experiencing the adverse health profile in question (as part of the status quo "health trajectory" that the individual's portfolio of current and future health states might include). Choosing to participate in the ongoing program will reduce the probability of experiencing this particular adverse health profile, while implicitly leaving the others unchanged.
>
> The "marginal utility of a year in ill health" is now interpreted like the marginal value of an attribute in a hedonic model. Number of years in ill health (probabilistically) is just one attribute of a particular illness profile.
>
> The confusing footnote 32 on p. 20 has been deleted, and the discussion about "fates worse than death" has been incorporated into the text along with the discussion of the interaction term between discounted sick-years and discounted lost life-years.

As a minor follow-on, how is nominal life expectancy calculated? Is it based on standard life tables and the respondent's age, sex, race, other variables? This does not seem to be reported.

> REPLY: Nominal life expectancy is based on an adjustment to the standard life tables. In focus groups and pretesting, individuals consistently overestimated their life expectancy. They talked about the age at which their longest-living grandparents or parents died. They tended to leave out of their calculations anyone who "died young" from something other than "old age." In response to this optimism bias, we presented them with their nominal expectancy as predicted by standard life tables plus a specified number of years. This greatly reduced widespread protests that they expected live longer

than what the life tables say. We differentiated between males and females by adding 7 years for men, and 9 years for women, to the average population-wide actuarial lifespan numbers.

Our survey does later ask people to report the age until which they expect it is most likely they will live. Fortunately, on average, for each age, these life expectancies line up reasonably well with the longer-than-actuarial lifespans we used in our stylized illness profiles. Future research will assess any biases from scenario adjustments with respect to subjective life expectancies, to the extent that they differ from the nominal life expectancies stated for each individual (based on their know age and gender) in our survey instrument.

On p. 16, it is stated that pdvp = pdve + pdvr. This suggests that respondents assume they do not pay for the health program during any period when they are ill with the illness that the program addresses, but resume paying for it once they recover (or experience remission). I do not believe [that?] this point, or what respondents were told about it, is specified in the text. Also, it does not seem very plausible that one would continue a program of diagnostic testing, etc., after recovering (going into remission) from the target illness, since some more appropriate follow-on treatment for those who recover is likely to exist.

REPLY: Let's take your concerns in turn. First, on Form 13 of the survey, we specify "If you have already suffered from one of these illnesses, please view these as possible recurrences."

We also tell respondents on Form 21 that "To make it easier to compare, we present all costs as monthly costs, and also as annual costs. You would need to pay for, and participate in, a program for the next 34 years to get its benefits." (Here, 34 years was the remaining nominal life expectancy for this particular individual.) Later in the instructions, individuals are led to understand that they would only have to pay for such testing while they are healthy or after they have fully recovered from an illness.

I do not understand the discussion of "mitigating bracketing biases associated with omitted substitutes" (p. 12). 1 think the idea is that in considering their WTP for one health program respondents should consider the other health risks they face and other methods they might use to reduce risk of the illness in question and others. Is the idea that by asking respondents about programs to reduce 10 different illnesses (i.e., two illnesses in each of the five conjoint choice tasks) over the course of the survey that they will be encouraged to recognize these broader issues? That seems somewhat in contradiction with the instruction to respondents to treat each choice independently, and of course it is not clear how this would affects responses to the first conjoint choice task. Also, Read et a1 1999 which is cited here is missing from the reference list so I could not turn to that for clues.

REPLY: You are correct that bracketing bias arise when individuals fail to consider the full range of substitute goods in their choice set and we wish to reduce this bias by bring to mind other significant health risks and then having respondents reflect on their subjective probability of each.

The goal of asking respondent to treat each if their decisions as independent of the others is not to discourage them from thinking about substitute programs. Rather we wanted to discourage them from framing their choices as a decision to purchase a package of programs. If their decisions cannot be treated as independent, we would have to condition each choice on all preceding choices in the survey. We wanted respondents to approach each choice afresh, as if it was the only one they were currently being asked to make.

The selectivity correction (pp. 21 -22) and associated model 5 is not well explained, and perhaps cannot be explained in a short-enough space to fit in this paper (which is already rather long). Since selectivity correction apparently makes little difference to the results, perhaps model 5 could be deleted and the authors could say they tried a model with selectivity corrections which had little effect.

REPLY: We have made room for a slightly expanded explanation since we would prefer to retain the selectivity correction given that it is statistically significant.

Empirical results

I find many of the empirical estimates to be implausible. While the estimates of VSL (and its relationship to age) are perfectly consistent with much of what is in the literature, the estimates of WTP to reduce morbidity seem incredibly large. For example, model **3** (Table 2) seems to imply that: (a) a year when ill is worse than a year dead (this might be true if the symptoms are severe, but the paper provides almost no information about symptoms); (b) a year when recovered or in remission is almost equally as bad as a year when dead (not at all plausible). Using the authors' preferred model, the disutility of one year of non-fatal illness is more than half the disutility of death, which also seems implausible (unless the illness is very severe). It would be interesting to report if the respondents were asked any risk-risk tradeoffs in which they could explicitly report how much risk of one illness profile they would trade for a change in risk of another.

REPLY: The heterogeneity across degrees of morbidity needs more explanation than we can give within the page limits of this paper. So we have decided to focus more on the mortality results, using the durations of morbidity and recovery/remission as crucial control variables.

That said, we wish to make several points about our morbidity and remission-recovery control variables. First, our illness profiles describe very serious illnesses. As we noted above, we were very careful to describe the severity of the expected pain and disability as well as whether major or minor surgery and/or hospitalization would likely be required. Second, we know of no other models in the literature which so comprehensively integrate both mortality and morbidity risk reductions for a range of major illnesses. Thus a direct comparison is hard to make with existing literature. Morbidity does matter. Our focus groups revealed that, as individuals age, they do begin to recognize that certain kinds of morbidity represent “fates worse than death,” and this is borne out by the coefficients on our interaction terms. Third, when comparing WTP for discounted lost life-years and sick

years, lost life years are always discounted more heavily than sick-years or recovered/remission years because they happen later. Fourth, we too were surprised by the recovered/remission-years results. We did not instruct people to pretend they would be “as good as new” after suffering a major disease such as cancer or a heart attack or respiratory disease. This is one area of the survey where additional instruction would have been helpful, since we might have tried to force people to make the assumption, however improbable, that they would recover one hundred percent from these major illnesses.

The "panel" structure of the data (multiple choices from each respondent) has the potential for confounding within and between respondent differences. As one example, I am unsure how to interpret the finding that the disutility of additional lost life years shrinks as number of life years lost increases. If this were true for a respondent, it would imply that marginal WTP to reduce mortality risk decreases as the age of death decreases, which does not seem very likely to me. Alternatively, if it is driven by differences between respondents, it suggests WTP to reduce mortality risk (holding latency of death constant) increases with youthfulness of respondent, but at a decreasing rate, which seems more plausible. Perhaps some sort of individual fixed effect would help address these concerns?

REPLY: We make use of the panel nature of the data and estimate all of our models using conditional logit algorithms which include individual fixed effects. Thus we believe that we avoid the problem of confounding “within” and “between” respondent differences. While the titles of the original Tables 2 and 6 mention this fact, we neglected to point this out in the original Table 3 or in the body of the paper. In any event, panel methods make a minimal difference because the characteristics of each illness profile were extensively randomized, independent from the respondent’s characteristics other than the fact that their current age limits the aggregate number of years in all three possible future health states. The randomized design of the illness profiles limits the potential for omitted variables bias of the sort that is commonly assessed in panel data.

In this analysis, marginal utility is conditioned on the age of the respondent at the time of the survey, but it is not conditioned on the age that the respondent *will be* when the future health state will be experienced. In a separate paper, we differentiate between “age-now” and “age-at-health-state” to reveal two different and countervailing effects. This generalization, however, requires extensive model development. This separate paper, which takes this paper’s findings as a starting point, is also struggle to keep under 40 pages.

The key detail to keep in mind is that the estimated (dis)utilities are for a *discounted* health-state-years. More undiscounted years, occurring far in the future, may convey the same utility now as fewer undiscounted years, occurring sooner.

Finally, why does the sample mean VSIP vary so much across models in Table 3? I would have anticipated that prediction at the sample mean would be much more stable.

REPLY: These are not fitted VSIP estimates at the means of the data, but the mean across the sample of the fitted VSIP values, calculated at the individual-specific vectors of explanatory variables for each observation. The manner in which a particular specification accommodates the tails of the distribution (typically very young and very old subjects), has a lot to do with the appearance or non-appearance of extreme values in the fitted point estimates of WTP for each individual, for each stylized illness profile used in the survey instruments people saw.

One of the main considerations in selecting an appropriate specification is the plausibility of the point estimates of WTP for individuals with extreme characteristics (especially age) and for illness profiles with extreme attributes (especially large numbers of sick-years or lost life-years, typically for the youngest respondents). Across the 15040 illness profiles used in the survey, none of the fitted individual WTP amounts should fail the "laugh test." Thus we added generality to the model as long as it was statistically significant and eliminated the anomalously huge or tiny WTP estimates that seem merely to be artifacts of a too-restrictive functional form.

We wished to see a plausible degree of variability in individual- and illness-profile-specific point estimates of WTP, but the distribution of illness profiles in the survey does not reflect the distribution of illness profiles in the real world, so these "sample mean fitted VSIP" statistics are not realistic either. Thus we decided to drop the "sample mean fitted VSIP" statistics and concentrate the reader's attention on the implied WTP estimates for our benchmark cases.

Minor comments

The paper claims that previous approaches to provide per-year health values like QALYs and VSLY do not take a utility-theoretic approach. I believe this is incorrect (or at least exaggerated). The VSLY idea is clearly that marginal utility per year of life is constant compared with a year of being dead, and there are explicit utility-theoretic justifications for QALYs, cited in Gold et al. 1996, for example.

REPLY: We agree with you with respect to QALYs and have tried to make this more evident in the text. However, the way that values for a statistical life-*year* (VSLYs) have been constructed seems far from congruent with standard marginal economic analysis. For a fuller discussion of VSLYs we now refer to J.K. Hammitt, "Valuing Changes in Mortality Risk: Lives Saved vs. Life Years Saved," *Review of Environmental Economics and Policy* 1: 228-240, 2007.

Since the interventions are described as including "diagnostic screening, remedial medications, and life-style changes" the valuation measures should be interpreted as net of any disutility associated with the intervention. In particular, respondents might perceive significant disutility associated with life-style changes and perhaps medications.

REPLY: Yes, we agree and considered this issue carefully during the development of the survey. We contemplated wording to explain the opportunity cost of time to get to the

doctor to have the blood test done, as well as the monetized disutility of the time and effort needed to cooperate with the doctor's instructions about medications and life-style changes. To preserve other questions and clarifications which were at least as important as this one, these details were eventually sacrificed. We faced a binding constraint to keep the survey at an acceptable length for our contract with the survey research firm. The reviewer's point is well taken, however. We acknowledge this issue at the end of the third paragraph at the beginning of the theory section.

7.2.2 Second round of reviews from the *AER*, with our replies

As mentioned above, we prepared and submitted our replies to each of the comments offered below, but Referee III abdicated from the review process, citing a conflict, without conveying any information about what he thought about our responses. Referee I did not reply and Referee II was not consulted further. Thus we have no idea which, if any, of the points raised in the first three reviews were not adequately dealt with in our responses below.

Note that the separate Appendices mentioned in the following section have now been incorporated into this document.

7.2.2.1 Editor's comments

On Referee I's comments, "I think you should work much more on the writing and exposition." I probably disagree with the referee to some extent, because I think I see the value-added of the paper pretty clearly. But the paper is written in a style to address specialists in the area and their concerns, not in a style for a more general audience. On the other hand, I understand that you need to provide a great deal of detail to specialists and convince skeptics about your data.

> *Response*: We are grateful that you see the value-added in the paper and that you appreciate our need to achieve a balance between writing for a general audience and writing to satisfy specialists who are more detail-driven. We hope that our current round of edits is helpful in this regard.

One very important change I recommend you make is to have an extensive set of Appendices, where you can put more detail about your survey, data collection, and estimation method than you have now. That would allow you to allude only briefly to issues in the text, perhaps in a summary form, and then refer the reader to the Appendix for the real details. An example of this is your quality-assurance discussion on pp.9-10.

> *Response*: Thank you for this suggestion. We have now developed six appendices, into which we have moved some manuscript material and through which we address the newer concerns of both you and Referee III and, to a lesser extent, Referee I. These appendices include:
>
> Appendix A: *Survey Design & Development*
> Appendix B: *Stated Preference Quality Assurance and Quality Control Checks*
> Appendix C: *Details of the Choice Set Design*
> Appendix D: *The Knowledge Networks Panel and Sample Selection Corrections*

Appendix E: *Model Estimation and Alternative Analyses*
Appendix F: *Estimating Sample Codebook*
UPDATE: All of these Appendices have since been consolidated into the present document (the Study Handbook).

However, I think your discussion of the modules prior to that is important and should be kept in the text, though it might be able to be reduced a bit. But I agree that the Introduction is too long and should be greatly shortened and written more succinctly. I also agree that you need to say something upfront about critiques of stated preference methods in general.

Response: We have shortened the introduction, while at the same time stating more directly the types of things our model allows researchers to do that they could not do previously. We also state explicitly why this capability is important to both researchers and policymakers. We also now address criticisms of stated-preference methods in the introduction, where we first mention that we rely upon a stated-choice survey, and in Appendix B (*Stated Preference Quality Assurance and Quality Control Checks*).

On Referee III's comments, I will leave your defense on why you think the subjects could answer the questions to you. Keep in mind that, if you put much of this in the paper rather than in a reply to the referee, most of it should go into an Appendix.

Response: We very thoroughly respond to Reviewer's III's central concerns in Appendices B and C, as well as to various other concerns in Appendices A and D.

I do agree with the referee that the paper is long and that the notation is complex. My own comments are as follows. First, I want to see a much more detailed Appendix on what the scenarios and profiles were that were offered to the subjects.

Response: We have shortened the paper slightly. We have tried to simplify the notation. To address the need for more detail, we now include variable definitions in the *Notes* that accompany our table of estimation results, and provide further details on the derivations of our formulas in Appendix E (*Model, Estimation and Alternative Analyses*). We also provide much more detail about the survey in our new Appendix A (*Survey Design & Development*) which describes the rationale behind the form and placement of every question in the survey. Appendix A includes (and refers to) one instance of the randomized questionnaire. Finally, we now provide a great deal more information about the design of the choice sets in Appendix C (*Details of the Choice Set Design*). This appendix explains the attribute ranges and algorithms that we used to generate each of the 7,520 essentially unique choice sets used in our survey.

Second, I want an Appendix on Knowledge Networks and more in the text on it.

Response: We now provide Appendix D (*The Knowledge Networks Panel and Sample Selection Corrections*). This appendix explains why we chose Knowledge Networks Inc over other survey firms—because of how they recruit and maintain their panel of respondents. In the second half of this appendix we explain how we evaluate and control

for possible sample selection bias in the estimates based on our usable sample. In the main footnote in the text that deals with Knowledge Networks, we now go into more detail as well.

Third, as an applied econometrician, I want to see the exact likelihood function that you estimated, for the three-choice model in an Appendix, and I think you need a much clearer discussion of your estimation in the text. The notation starts to become cumbersome in equations (3)-(4), but I was looking right after that for a logit model formulation of the problem, but that never really came. Later, you said it is a "fixed effects" logit (what are the fixed effects?). If it's a fixed effect across people, then usually one cannot estimate intercepts because they are not identified. You need a real discussion of estimation before you start discussing using the estimates to calculate *VSL*s or anything else.

Response: We now try to do a better job of this, and to simplify the notation as you and Referee III requested. In addition, we have provided a greatly extended description of the model and of the estimation method in our new Appendix E (*Model, Estimation and Alternative Analyses*). This appendix provides plenty of detail concerning the fixed effects logit estimator used to produce the vector of maximum likelihood parameter point estimates and the parameter asymptotic variance-covariance matrix reported in the paper (both the biostatistical perspective from the Stata manual and the econometric perspective from Greene's standard graduate textbook).

Details: The first notable thing about the structure of our data on respondents' three-way multiple discrete choices is that these are effectively "panel" data. Each respondent, typically, provides us with five different choices. With panel data, there is always a question whether a set of slope coefficients, estimated using simply the pooled data without recognition of its panel nature, might be affected by heterogeneity bias. (Heterogeneity bias is a form of omitted variables bias, where the explicit explanatory variables are correlated with unobserved forms of heterogeneity across individuals, so that the estimated slope coefficients are biased). Fortunately, the randomized design of all of our choice sets, conditional only on the age and gender of the respondent and the plausibility of some types of outcomes, means that the x_{ij} variables in our models are unlikely to be correlated with any omitted variables, especially since we control for the respondent's current age in our models.

Nevertheless, the fact that we have repeated choices for each person in our sample immediately led us (and almost every other reviewer of our work) to a concern that appropriate panel-oriented econometric methods should be used with these data. The parameters of our model are thus estimated using the fixed effects conditional logit choice model as implemented in the Stata 10 econometric software package. The model is described in considerable detail in the Stata 10 Reference Manual under the heading "clogit – Conditional (fixed effects) logistic regression" (p. 285-287). We also now provide two versions of the rationale for this model in Appendix E, one from the biostatistical tradition of analysis and one from the econometric tradition of analysis.

In Appendix E, Section 4.2 (Econometric Perspective), we provide in equation (20) the unconditional likelihood function underlying our estimates, in the case of three alternatives per choice and five choices, with the link to the variables in the paper explained in the preamble to that equation. The parameters are actually estimated using a

conditional likelihood that nets out the fixed effects, as shown in equations (21) through (24) of Appendix E, although we revert to the simpler two-alternative, two-choice case to keep the notation manageable. We also include a discussion of the appropriate Hausman text to use in this context, and the results of this test for our sample.

I also wonder about the root function for income; why not just have a polynomial or some flexible-form specification with more parameters? I have to admit, I stopped reading on pp.16-17 because I didn't understand the estimation.

Response: First, we have sought to clarify and simplify our estimation so that it is more understandable.

Second, given the vastly greater convenience of a fixed transformation parameter in terms of the estimation, we elect to approximate preferences using this particular transformation, which is close to a square root function. While this function is less flexible than a quadratic form in net income, it allows for risk aversion with respect to net income but still guarantees monotonicity, which is also desirable. A polynomial (quadratic) form allows for a non-monotonic relationship and high-income outliers can display negative marginal utilities of income simply as an artifact of the best-fitting quadratic form on the increasing side of the function (where most of the mass of the data resides).

While we might accept a marginal utility of income close to zero at high incomes, we are unwilling to accept *negative* marginal utility at high incomes, especially when we suspect this is merely an artifact of the best fit in another part of the net income domain. Since the marginal utility of income resides in the denominator of the *WTP* formula, it is very *inconvenient* to have this derivative pass through zero into the negative range. As it does so, the implied *WTP* first gets extremely large and positive, then it jumps discontinuously to a very large negative number. Monotonic forms prevent this outcome.

A detailed question I have is: did you take into account the impact of future illness on work and reductions in earned income, which would feed back into Y?

Response: Yes, it is possible in our model to assume different things about how individuals expect their future incomes to change if they get sick. In the main paper, for the sake of simplicity, and in keeping with lifecycle models of income, we assume that individuals believe they will be able to smooth their income even in periods of illness. We mention this in the first paragraph of Section II.C. If you would prefer a different set of assumptions about future income, our model does allows the researcher to make whatever assumptions they wish about what individuals typically anticipate. We now include, in Section 6 of Appendix E (*Model, Estimation and Alternative Analyses*), a set of results generated under the assumption that income will fall by half if the individual gets sick with the illness or injury in question. However, since discounted sick-years are often small relative to total discounted remaining life-years, this adjustment actually makes only a minor difference in the estimated parameters and *WTP* results.

Finally, I would ask that you use full double-spacing of the paper, with the same font you have used, but you could reduce the top, bottom, and side margins to 1-inch if you like.

Response: Yes, easily done. We have also added a considerable number of sub-headings and sub-sub-headings to add more structure to the paper. This has stretched out the space a little, but has made it easier to read. We have eliminated two tables and trimmed quite a lot of material. The text of the paper is now around 36 pages long, but with its numerous references, two tables, three figures and a one-page appendix, it extends to 45 pages. We would be happy to try to make a further round of cuts if this would be appropriate, but we would appreciate some guidance on what else can go, given that the referees wished to see additional discussions.

It has been somewhat difficult to respond to all of the considerations raised by the referees without lengthening the paper too much, however, the paper is now supported by five extensive appendices, intended for online posting. These appendices include almost 140 pages (single-spaced) of additional discussion and alternative specifications. This page count does not include the single "instance" of our randomized survey instrument included with Appendix A (*Survey Design & Development*). As the extensive questions and comments of Referee III indicate, it has been difficult to pare down the wealth of issues involved in this study so that they fit into a single succinct paper.

We apologize that our responses to the editor's and referees' comments contained in this document now extend to 36 pages (but single-spaced in this case). We hope that the level of engagement of our diligent referees translates to the *AER*'s wider audience.

7.2.2.2 Referee I's Comments

General Comments

The introduction goes from page 2 to page 5. It is difficult for the reader to learn the main issues being addressed by this paper, its main value added, its main findings, and the major implications of the findings from the introduction as written. Why specifically is this paper of interest to a general audience of economists? There is a discussion of value-added on page 27. More generally, the paper, in its current version, reads more like a report on a contract than like a journal article. As written, it may be appropriate for a specialty journal such as *Journal of Public Economics* or *Journal of Risk and Uncertainty* or *Journal of Health Economics*, where more readers know this literature, than for a general interest journal. Even then, readers of a journal like the *JHE* will want more detail such as that for which I ask below. If rewritten for a general interest journal, the paper needs to be more focused. The value-added needs to be much clearer to the general interest reader.

Response: We now state more directly those things that our model allows researchers to do, that they previously could not do, and why this is important. In the process, we also state explicitly the value-added and major implications of our proposed approach. In addition, we have shortened the introduction and removed some of the detail. Your suggestions have pushed us to make the introduction much more relevant and engaging for the general reader. Thank you.

But it is the meaning of latency and severe prior morbidity that is not adequately described in the paper.

Response: We now define "latency" when it is first used. The latency period of an illness for an individual is the period of time in years between the present and the probable onset of a given illness. "Prior morbidity" occurs when the individual has previously experienced or is still experiencing a major illness such as the ones we evaluate. However, to shorten the paper we have cut that section and we no longer refer to that term in the text.

Tables need better explanations.

Response: Throughout the paper now we take great care to clearly and thoroughly explain each of the tables. The new Table 1 is described in Section III.A. Tables 2, 3, and 4 are described in Section III.B. See also the new and more-detailed *Notes* accompanying each table itself. Thank you for this suggestion.

Specific Comments

p. 2. Illness trajectories over the life cycle are often not even known to researchers. Is this not overpromising?

Response: Illness profiles are presented to respondents as probabilistic events so they only need to be *plausible* to respondents. Test subjects understood that the baseline probabilities we presented to them were to be taken as "best guesses" about their future illness risks, given their current age and gender. We now make it clearer to the reader in Section I.A.(ii) that the illness trajectories (profiles) described to each respondent are *hypothetical* and meant to be interpreted as the central tendency for an individual of that age and gender, for the named risk that is to be reduced by the program in question. It proved logistically infeasible to communicate to survey respondents a distribution of possible illness profiles.

Future survey researchers may invent ways to convey a joint distribution of time periods in each future health state, and redacted medical records may in the future be available so that patient confidentiality does not prevent us from knowing the true joint distribution of time in each health state for particular types of diseases. But these two needs cannot be satisfied at present. For now, we only expect respondents to understand the illness profile we describe as a plausible one for them. We paid close attention to the plausibility of different illness profiles for the different names of the diseases. For example, as shown in Table C1 in Appendix C (Details of *the Choice Set Design*), nobody dies suddenly from diabetes or Alzheimer's disease.

p. 2. Under the second point, research on competing risk of death should be cited.

Response: Yes, thank you. In an earlier version we cited the Dow et al (1999) paper when referring to research on competing risks of death. We have reinstated this reference in response to this suggestion. Are there additional papers that this referee would recommend?

p.2. Last line. Who are the exceptions to this rule? Some citation is needed to back up this generalization.

> *Response*: We now report where other researchers have explored the effect of specific factors on *VSL* estimates such as age, income, health states, and latency. See the footnotes to our "first," "second," and "third" points on pages 3-4.
>
> For example we note that researchers have explored the influence of each of these factors on *VSL*s but not in a comprehensive structural model of inter-temporal demand. For age, see Krupnick (2007) and Viscusi and Joseph E. Aldy (2007). For income, see Janusz R. Mrozek and Laura O. Taylor (2002), Viscusi and Aldy (2003), and Dora L. Costa and Matthew E. Kahn (2004). For future health states, see Krupnick et al.(2002) and Alberini et al. (2004).
>
> We also point out that other researchers have valued risk reductions at one time in the future (e.g. Alan Krupnick et al. 2002, Anna Alberini et al. 2004, James K. Hammitt and Jin-Tan Liu 2004, and George Van Houtven, Melonie B. Sullivan, and Chris Dockins, 2008) but not the reduction of risks involving time patterns of several different adverse health states.
>
> With respect to research on latency we note that Van Houtven, Sullivan, and Dockins (2008) use a survey that asks respondents to consider a forced relocation, for one year, to one of two other cities, where the two locations differ only in their relative and absolute frequencies of fatal stomach, liver, or brain cancer versus car accident deaths. They randomly describe the illness profiles for the cancer as having 5, 15, or 25 years of latency and either 2 or 5 years of morbidity.

p. 3. Perhaps policy evaluations use the same *VSL* for persons of all ages and for the rich and poor, but researchers certainly have addressed this issue. Also some policy evaluations work in life years and value of life years rather than in *VSL*.

> *Response*: Yes, we agree with the first assertion. Researchers such as the ones we note in the preceding response have explored the effects of age and income on estimates of the *VSL*. We now make the contributions of these researchers clearer in the text.
>
> But this referee also raises the issue about policy evaluations using the value of a statistical life-year (*VSLY*). Our broader point is that the current construction (and use) of *VSLY*s is likely to be both inaccurate and largely invalid because it is commonly constructed by dividing the measure of a one-size-fits-all *VSL* by the average remaining lifespan, although we now note the contribution of Moore and Viscusi in a footnote to the third paragraph of our introduction. Our paper offers a utility-theoretic approach to the construction of values for individual prospective years in adverse future health states. We agree that some policymakers such as Graham (2003) have recommended using both *VSLY* and *VSL*. Our hope is to improve the empirical methods used to calculate *VSLY*-type measures.

p. 3. The public's semantic confusion with *VSL* is a reason to improve risk communication, not as a rationale for a particular analytic approach. There are better reasons, such as, the range in which risk is typically evaluated is closer to micro risk.

Response: Yes, we agree. We have removed our extensive discussion of the rhetorical confusion often surrounding *VSL*s. Instead, we emphasize the increased validity and accuracy of our approach to valuing micro-risks, and the relevance of this new approach (and alternative normalization) to both researchers and the policymakers who use these estimates. We now offer only a brief acknowledgement of the common semantic confusion.

p. 3. There have been critiques of stated preference approaches. Should not the reader be aware of these and your take on how you have accounted for these criticisms in your work if you have? I am not sure this goes in the introduction, but it goes somewhere.

Response: Yes, we note these critiques and concerns as well as our efforts to address them. We footnote some of the controversy concerning stated preferences where we first mention that we use this method for this study. We now go into much more detail concerning specific biases and our efforts to evaluate and avert them in Appendices A (*Survey Design and Development*) and B (*Stated Preference Quality Assurance and Quality Control Checks*).

For example, we discuss how stated preference methods generated controversy in the past because of concerns that people would overstate their willingness to pay for a public risk reduction. However, over the past ten years, important strides have been made in understanding and minimizing concerns about the incentive compatibility of these choice situations (John List 2001). Indeed a recent meta-analysis shows that stated preference estimates of the *VSL* are systematically lower than those produced by revealed preference data (Ikuho Kochi, Bryan Hubbell, and Randall Kramer 2003). A variety of additional validity checks are covered in Appendix B (*Stated Preference quality Assurance and Quality Control Checks*).

p. 4. In general, there is too much detail here for an introduction. The reader who is deciding whether or not to read this paper gets mired into details of the conduct of this study.

Response: We agree. As mentioned above, we have rewritten the introduction to sharpen the reader's focus on our contribution and its implications. We have reduced the length of the introduction to just thirteen streamlined paragraphs.

p. 4. bottom. What is meant by "program choices"?"

Response: In our stated-preference choice scenarios, individuals are offered illness-specific risk-reducing "programs" which involve annual diagnostic testing and, if needed, remedial drug therapies and instructions on related lifestyle changes that the individual should undertake. Each of these programs has a different annual fee as associated with it. We make this terminology clearer now, when we first introduce the term.

p.5. This page deals with the mechanics of the study. What about what we learn from this analysis?

Response: Thank you for pointing out this oversight. We have substantially re-written the introduction to make the value added from our study, and lessons learned, more clear from the outset.

p. 6. State what you did first and then discuss positive attributes and perhaps weaknesses of your survey. You say the survey is representative but don't really describe the sampling frame, characteristics of participants, etc.

Response: We have substantially reorganized this section in response to this request and the recommendation of the editor. We now include six comprehensive appendices which contain much of the prior material from this section of the paper plus other important background information. The editor requested that we leave in a description of the survey modules. However we have created some new appendices which respond to this point—Appendix B (*Stated Preference Quality Assurance and Quality Control Checks*), Appendix C (*Details of the Choice Set Design*) , and Appendix D (*The Knowledge Networks Panel and Sample Selection Corrections*). These changes seem consistent with this referee's preferences too.

p.7. Is the course of these diseases really known? What about a heterogeneous category such as heart disease? What is the source of information from which information on the trajectories is drawn?

Response: We addressed this issue earlier in our responses. These illness profiles are not presented as "known" or "certain" but as possible (probabilistic) health events. These are hypothetical illness profiles that we expect respondents to treat these as "typical" for someone of their age and gender. A great deal of interviewing and pre-testing was conducted to ensure that respondents would find these randomized elements of the illness profiles reasonable and plausible. We now state explicitly in the paper that the illness profiles and program costs are hypothetical.

Table 1 is hard to interpret. The reader needs more help with this. How are pre-illness years defined? Where does the data on monthly cost come from? What is included in monthly cost? In general, we need to know how these values were calculated.

To shorten the paper, the old Table 1 has been moved to Appendix C (*Details of the Choice Set Design*), where it is now listed as Table C1. In the paper, we now make it clear on page 10 that each health profile is constructed with extensive randomization, for each individual (given their current age and gender) and for each disease (subject to constraints on what is plausible for that illness).

Ten of eleven possible illnesses were selected for each survey instrument and ten randomized and unique illness profiles were generated in advance for each panelist who would be invited to participate in the survey. For example, no female respondents were asked about prostate cancer, and no male respondents were asked about breast cancer. Likewise, nobody was asked about an illness profile where the illness would strike them at an age younger than they were at present (or indeed, within two years of the present, to ensure plausibility).

The number of pre-illness years (the same thing as the "latency," the time period before symptoms begin to appear) is varied from person to person and from illness to illness. This is necessary for us to be able to identify the effects of latency on *WTP* for a health risk reduction.

We now make it clear that the stated costs are hypothetical.

pp. 6 ff. Why does the reader need all of this detail on modules? Typically, papers do not describe survey formats in detail. This is a lot of detail, in particular, when detail on the tables is lacking.

Response: Except for the requests made by the editor, we have moved much of our description of the actual survey, and one complete example, to Appendix A (*Survey Design & Development*). At this referee's request, we provide a much more detailed description of the tables now, both in the text and especially in the *Notes* accompanying each table.

p. 11. Give an operational definition of the latency period. How is a starting date assigned to the latency period? What are respondents told about this? Recovered/remission states typically involve a time period, and this varies among diseases.

Response: Latency periods (time until symptoms would begin) are described in terms of the age the person would be when the illness started (always an age at least two years into the future). Recovered/remission states indeed involve a time period, and these time periods vary across the diseases. Like each other distinct time period, the recovered/remission period is randomized in terms of its length. For each illness profile, the individual's remaining lifetime is divided into four time periods, any three of which can be randomized and the fourth defined by their nominal life expectancy. These randomized intervals are converted into verbal descriptions for the choice scenarios in our survey.

p. 18. The reader needs more help with Table 2.

Response: The old Table 2 displayed the results from a selection of preliminary models, designed to show the robustness of the key variables. To shorten the paper, we have moved these preliminary models to Appendix E (*Model, Estimation and Alternative Analyses*). Section 6.1 of this appendix discusses these models, and the old Table 2 now appears this appendix as Table E1. The first paragraph of section III in the paper now describes these preliminary models briefly, and refers the reader to Appendix E for details.

p. 19. This seems to be the first time we are told about the 500,000 Knowledge Networks sample. The reader is told next to nothing about Knowledge Networks. How the sample is drawn, biases, etc.

Response: We agree that we needed to offer more detail. We now mention Knowledge Networks in the first paragraph of Section I (*Survey Methods and Data*) of the paper and

explicitly refer readers to Appendix D (*The Knowledge Networks Panel and Sample Selection Corrections*).

p. 19. While a 6-point improvement in the likelihood function is something, it is not much. Again as with previous tables, readers could use help in interpreting the results in Table 3.

Response: The old Table 3 is our new Table 1. We provide a fuller description of the new Table 1 in the text, and have augmented the *Notes* to the table to include variable definitions and other details.

p. 24. top. The paper goes over a controversy about the income elasticity of *VSL* from various papers. It is not helpful in helping the reader know what the issues are. Is the Hall and Jones framework really sufficiently similar to the others for a comparison to be made?

Response: We mention that the income elasticity is an issue because this is the one dimension along which policy-makers have demonstrated a willingness to adjust *VSL*s over time (to reflecting rising real incomes). We have reframed our discussion of this topic to make it more relevant, and we have demoted the point about health insurance, raised by Hall and Jones, to a footnote to section III.B(iii).

p. 25. Is the latency period from the point of exposure to a harmful substance or engaging in a harmful activity and the occurrence of an adverse health effect? Then what is the latency period for colon cancer?

Response: The latency period is the spell of time in the current health state before the illness or injury would produce symptoms, should the individual succumb to this illness or injury. The latency period ends when pain and disability due to the specified illness would begin. Most people assume that this latency period is determined by factors such as their age, gender, genetics, and other personal characteristics as well as environmental exposures.

The comparisons with other studies are potentially interesting, but so much territory is covered that nothing is really described sufficiently well.

Response: We have streamlined this section, and have added more structure in the form of subsection headings to delineate each category of results and the corresponding insights from the previous literature. While it has been a challenge to "describe better" while not lengthening the paper, we hope that our changes are helpful.

p. 27. The notion of "allocating risk-reduction expenditures across health risks" is not adequately described in the paper.

Response: This phrase previously appeared in a passage that has now been rewritten to form the first sentence in Section I (*Survey Methods and Data*) in the revised paper. The replacement text reads: "It is very difficult to identify market data that would adequately illustrate differences in individuals' demands for reductions in the wide variety of health

risks that may come to bear across their remaining years of life." The expression was also used in our concluding section, where it has also been changed. The passage "allocating risk-reduction expenditures across health risks" no longer appears in the paper. Thank you for helping us to express this more clearly.

7.2.2.3 Referee II's Comments

I think this is a nice paper and much improved from the earlier version I reviewed. The survey from which the data come is nicely constructed and the models analyzed in this paper seem like a plausible first (or second) order approximation to how *WTP* may vary with risk of illness.

Response: We are grateful that this referee seemed to be pleased with our first round of revisions.

At this point, I have only a few minor comments to offer.

p. 1 (and p. 21, fn 46). While I agree that many estimates of *VSLY* are obtained by dividing an average *VSL* by an average life expectancy (perhaps discounted), that is not a fair characterization of Moore & Viscusi 1988.

Response: We agree and have rewritten the introduction and included a footnote to clarify that Moore and Viscusi (1988) take a different approach.

p. 2. I am confused about the construct validity comments. The first paragraph says "the measure varies systematically," which sounds like it means that estimated *WTP* varies systematically (in the anticipated direction with health attributes). The second paragraph sounds like it is describing the way in which the health risks are described to respondents "By more completely defining the difference …" Please clarify whether the claim here is that the survey asks about a better defined commodity than in some previous work, or that the results vary in accord with expectations.

Response: We are asserting both that the commodity being valued (health risk as an illness profile) is better defined than in previous work and that our results vary as expected. An additional advantage of this approach is that it incorporates more individual-specific information into the estimate of the value of a future lost life-year. In contrast, both the applications of the *VSL*s and *VSLY*s to latent risk reductions currently require the researcher to supply this information when making assumptions about the future actual value of avoiding a lost life year, substitution between future lost life years and future sick years, future income flows, risk preferences, discount rates, etc.

However, we have removed the language about construct validity in an effort to make the introduction shorter and more friendly to a general audience. We hope the earlier ambiguity has been eliminated in the course of our major revisions to the introduction.

pp. 2-3. I find the claim that researchers treat *VSL* as independent of income, life expectancy, etc. to be too strong. As is clear from many of the papers cited here, researchers know well that *VSL* varies with income and is likely to vary with life expectancy and other factors, and there is much

research investigating these topics. The complaint would be more appropriately addressed toward government use of *VSL* estimates in which these sorts of variations are typically ignored.

> *Response*: Yes, we agree. In the footnotes to the introduction, we now report where other researchers have explored the effect of specific factors on *VSL* estimates such as age, income, health states, and latency.
>
> For example, we note that researchers have explored the influence of each of these factors on *VSL*s but not in a comprehensive structural model of inter-temporal demand. For age, see Krupnick (2007) and Viscusi and Joseph E. Aldy (2007). For income, see Janusz R. Mrozek and Laura O. Taylor (2002), Viscusi and Aldy (2003), and Dora L. Costa and Matthew E. Kahn (2004). For future health states, see Krupnick et al.(2002) and Alberini et al. (2004).
>
> We also point out that other researchers have valued risk reductions at one time, or at a few different times, in the future (e.g. Alan Krupnick et al. 2002, Anna Alberini et al. 2004, James K. Hammitt and Jin-Tan Liu 2004, and George Van Houtven, Melonie B. Sullivan, and Chris Dockins, 2008) but not the reduction of risks involving essentially a continuum of time patterns of several different adverse health states.
>
> With respect to research on latency, we note that Van Houtven, Sullivan, and Dockins (2008) use a survey that asks respondents to consider a forced relocation, for one year, to one of two other cities, where the two locations differ only in their relative and absolute frequencies of fatal stomach, liver, or brain cancer versus car accident deaths. They randomly describe the illness profiles for the cancer as having 5, 15, or 25 years of latency and either 2 or 5 years of morbidity.

p. 10. I am surprised by the claim that, "as economic theory would predict," *WTP* "rises with the expected incidence of health risks in future years." The reference for this claim is suppressed for anonymity so I cannot tell, but is this paper published and available? Is it possible to give the intuition for the result here? Anticipated future health effects can influence both numerator and denominator of the *VSL* (or VSIP), so, in general, effects like expected future health and life expectancy have an ambiguous effect on *WTP* (see, e.g., Hammitt 2007).

> *Response*: Ehrlich (2000) predicts that as individuals age, their stock of health declines. This exogenous decline in the stock of health implies that the value of avoiding marginal reduction in health stock increases with age. We hasten to point out that, in Ehrlich's framework, this is only one element of the current of value of a future health-risk reduction. (The present value of remaining future consumption is another component (or set of components, since Ehrlich can distinguish the marginal utility of a future year of consumption which may also vary with age as it affects the number of years of remaining consumption).
>
> Our data actually *can* show that the marginal value of reducing a given risk increases with the future age at which the illness or injury would come to bear; in other words, individuals view it as optimal to allocate more money to protect the health of their older selves from further deterioration. However, the alternate model that produces this insight is not within the scope of the current paper. In that model, utility is modeled as depending separately on the individual's age at the time they are making their choice among policies (quadratically as we do here), and the age that they will be (in the future),

during each year of each adverse health under the specified illness profile (where these future ages also enter quadratically). This dependence of utility on future age complicates the model because it is then no longer the case that within health states, we can assume that all variable are invariant over time. Age will increase within each health state. Thus the algebra is rather untidy. This different model is the subject of a separate paper, "Two Types of Age Effects in the Demand for Reductions in Mortality Risk with Differing Latencies," by DeShazo and Cameron (2005), available online at http://www.uoregon.edu/~cameron/vita/VSL_age_120505.pdf . Rather than lengthening the current manuscript by digressing further along these lines, we have removed the assertion and refer to the Ehrlich model only in passing.

7.2.2.4 Referee III's Comments

This manuscript reports results of a stated-preference study with very ambitious goals. The authors aim to "estimate individual-specific schedules of expected [monetized] utilities for morbidity and mortality risk reductions in each year of an individual's remaining life," for twelve health risks over a range of risk reductions. With the exceptions noted below, the authors generally employ good-practice or best-practice methods for survey development and subject the data to the full range of skills of an accomplished econometrician. Nevertheless, for all its analytical sophistication and claims of success, not to mention length, the manuscript is likely to fall short of persuading many readers that subjects actually performed the very difficult preference elicitation task required to achieve the study objective.

Response: We are grateful for the huge amount of effort this referee clearly put into the task of reviewing our paper. It has been a long time since either of us has received such a detailed referee report.

Based on the opening remarks to these comments, this referee seems to be concerned primarily about whether our research subjects could actually handle the choice tasks we presented to them. We have tried to respond to this referee's concerns here, as well as to present these same concerns to the reader. While opportunities for this arise throughout the manuscript, we also devote extensive attention to these issues in our six new very detailed appendices, intended to be offered online. To review, these include:

Appendix A: *Survey Design & Development*
Appendix B: *Stated Preference Quality Assurance and Quality Control Checks*
Appendix C: *Details of the Choice Set Design*
Appendix D: *The Knowledge Networks Panel and Sample Selection Corrections*
Appendix E: *Model Estimation and Alternative Analyses*
Appendix F: *Estimating Sample Codebook*

Appendices A, B, and C, in particular, include a considerable amount of material that responds to this referee's concerns, itemized below.
UPDATE: As noted above, these appendices have now been incorporated into this Handbook

Validity of the Preference Data:

When given a difficult cognitive task with little incentive to devote much effort, rational subjects are likely to employ simplifying decision heuristics. These heuristics may provide considerable structure to the choice data. However, the data may not be informative about the underlying welfare-theoretic preference constructs of interest to the researchers.

> *Response*: Respondents' use of heuristics in decision making is indeed a very important consideration and one to which we devoted a great deal of care to minimize and evaluate. We now discuss the question of heuristics explicitly in Section 10 of Appendix B (*Stated Preference Quality Assurance and Quality Control Checks*), and less formally in several places in Appendix A (*Survey Design & Development*).
>
> An important question is whether SP data are more likely to be affected by heuristics than would comparable RP data. Our respondents probably see more information, more comparably presented, than they would be shown in any real choice situation with respect to opportunities to reduce risks to their lives and health. Moreover, we perhaps devoted more time and provided more learning opportunities (through our explanation of risk measures and use of risk graphics) to prepare them for their decision making than they would get during a typical doctor's office visit.
>
> The next concern is whether respondents selectively discard or recode the information presented in the choice scenarios in a way that renders their choice data unusable for the purposes of recovering an accurate estimate of their *WTP* for health risk reductions. As experimentalists, we acknowledge first of all that it is not possible to observe individuals' actual decision processes, so we cannot confirm or refute the presence of heuristics. All one can do is to look within the data for any evidence suggesting damaging consequences of heuristics.
>
> Do we see any blatantly obvious evidence of the use of damaging heuristics? No. First, the attributes of the illnesses and the characteristics of the respondents are strongly statistically significant. Second, the estimated marginal utilities and *WTP* measures vary as both general economic theory would predict and (for illness profiles where evidence exists) as most prior studies on health risk have found. Third, the relevant ranges of our final estimates are generally consistent with the available benchmarks for RP data that exist within the literature.

The authors anticipate this concern by acknowledging "the cognitive complexity associated with the choice task, which we seek to minimize through careful survey design, and which we evaluate carefully *ex post*." They assert that they conducted "a very wide array of robustness and validity checks." They list checks for risk comprehension, scope effects, order effects, scenario rejection, and sample selection biases, but fail to persuasively support this claim.

1. Risk comprehension. The authors administered "an extensive tutorial" on risk and probability, followed by a quiz to determine whether subjects could correctly answer "a simple risk comprehension question."
2. Sensitivity to scope was evaluated by verifying that "a lower cost and a greater risk reduction make a program more attractive."

3. The authors find little evidence of order effects, except in the last question in the task sequence.
4. The authors drop responses (but not subjects) where subjects chose the opt-out alternative and explained that the reason was "I did not believe the programs would work," indicating scenario rejection.
5. After each question they ask subjects directly how difficult the choice was on a scale of 1 to 5. The mean was 3.2.
6. Finally, the authors find that most of their estimates have the correct sign and good statistical significance.

The evidence offered on validity generally is at best indirect.

Response: While there are many distinctive threats to validity to guard against, this referee seems to focus upon and prioritize two particular threats: inconsistently evaluated tradeoffs, and the recoding of risk levels into simpler metrics. Because of the importance of this referee's concerns, we want to address them directly here before proceeding to the details.

Concerns about inconsistent tradeoffs. Your first concern is that, due to the complex nature of the choices, individuals may not correctly and consistently evaluate the risk-tradeoff questions. (This seems to encapsulate this referee's concern about items 1, 3, 4, and 5.) If there were pervasive problems, respondents' choices would not be internally consistent. For example, if the complexity overwhelmed respondents to a point that their choices did not preserve the property of transitivity, a degree of randomness would characterize the choice data. In the extreme, the observed choices would appear to be predominantly random, and our slope coefficients on the attributes of alternatives in the choice scenarios would be statistically indistinguishable from zero. Choice inconsistency thus does not appear to be happening in the extreme, and certainly not so much as to prevent us from getting reasonably precise measurements of a central tendency of each empirical preference parameter.

We acknowledge that some inconsistent choices may certainly occur in our data. (Some inconsistent choices are likely to be present even in many samples of RP data.) In light of this, we have looked into the possibility of employing a set of internal consistency checks provided in some software designed for evaluating conjoint choices in a non-parametric setting. Section 11 of Appendix B (*Stated Preference Quality Assurance and Quality Control Checks*) explains the types of validity checks offered in this software, but also details why our data appear not to be amenable for use with this software. These reasons are also itemized in the discussion below.

Recoding and Scope insensitivity. Your second concern is that even if choices are not random, respondents may recode the risk information, converting it from the metric in which it is described in the survey into one that is simpler to them but unobserved by the researcher. It was the possibility of such recoding, and the resulting insensitivity to the scope of good, that led researchers to demand scope tests in CV studies. We agree that the proper way to conduct the required "external scope test" is to "split the sample" into at least two groups and to administer different stated benefits in the choice scenarios presented to these different subsamples.

The referee is correct that if *all* respondents saw the same five choice sets, we would not have met the scope test requirement even if we had shown that *WTP* was larger when benefits were larger "between scenarios." This would have been the "weaker internal scope test" to which the referee refers.

In fact, we did undertake a strategy analogous to splitting the sample in the way we designed our choice sets. We apologize that we did not make it adequately clear in the last version of the paper that our choice scenarios vary across respondents, as well as across choice sets for each individual respondent. Since every choice set is uniquely randomized, we have actually pushed the "external scope test" strategy to its logical extreme.

Below, we will now respond point-by-point to this referee's individual concerns, in greater detail:

Retaining only subjects who could answer the risk question does not ensure that the remainder of the sample correctly and consistently evaluated the risk-tradeoff questions.

Response: Yes, this is true. A few respondents who didn't understand the risk questions may have guessed the right answer on the risk-comprehension question, just as a few of the people we excluded this way may have inadvertently selected the wrong answer despite being very comfortable with the concepts of risk that are employed in the survey. (To use a statistical analogy, there are probably a few "Type I" and a few "Type II" errors associated with this exclusion criterion.)

However, use of this "skill-testing question" does demonstrate our effort to avoid the inclusion of overtly inattentive or uncomprehending respondents who would likely produce inconsistent data. The counterfactual—not including, or not using, such a test—would have been seen by most researchers as a failure to assess and verify the likely risk comprehension of respondents.

The authors claim that, "even in the simplest possible choice models, individuals readily pass the scope test". However, the scope test does not satisfy the accepted standard for testing sensitivity to scope in stated-preference studies. The standard test requires splitting the sample, varying the size of the offered benefit between the two groups, and then checking for differences in valuations that are suitably large relative to the difference in benefits. At best, the authors' test is a much weaker "internal" scope test, i.e. whether the average subject noticed that the benefit was larger between scenarios.

Response: Our new Appendix C (*Details of the Choice Set Design*) now provides complete details on the range of possible values used for each attribute and the criteria for rejection of individual randomly generated combinations of illness profiles and risk-reduction programs. The choice scenarios offered to each individual were keyed to each person's gender and age, so we could not design, *ex ante*, a finite number of survey instruments with random assignment of these instruments to all respondents.

As outlined above, we did undertake a strategy that involves "splitting the sample" when we designed our choice sets for each individual. In the earlier version of the paper, we did not make it adequately that the choice scenarios vary across respondents, as well as across choice sets for a given respondent. For each of the 1,801

individuals whose choices were used to produce our estimates—indeed, for all 7,520 choices (i.e. for all 15,040 "programs") analyzed in our data--the illness profiles were likely to have been unique.

Instead of splitting our sample of 1,801 individuals into just two groups, and showing everyone within the same group the same set of choice scenarios, we effectively had 1,801 different groups of respondents, each containing just one person. We contend that this strategy actually vastly outdoes the usual "external scope test" because every respondent considered different illness profiles and different risk-reduction program costs.

Dropping observations based on self reports of scenario rejection for opt-out choices does not ensure that the same subjects or other subjects generally satisfied theoretical validity conditions for other questions.

Response: This is certainly true. Still, failing to drop observations where respondents directly admitted to pure scenario rejection would have been unacceptable (we are confident that this is not an alternative that this referee would prefer).

Our new Appendix D (*The Knowledge Networks Panel and Sample Selection Corrections*) now contains an extensive discussion of our *ex ante* criteria for rejecting certain respondents or certain choices from the estimating sample. (See sections 3.2.1 to 3.2.5 of this Appendix.) In the last paragraph of Section I of the paper (*Survey Methods and Data*), we now mention that "Appendix D gives the details concerning three ex ante criteria used to exclude certain respondents and/or choices from the estimating sample."

Self reporting of how difficult they thought the questions were is an unreliable measure of data quality. In a study using a similar self assessment, we found that subjects who reported that the questions were relatively difficult still performed well on utility-theoretic internal validity tests such as transitivity.

Response: Again we agree and believe that this should be an area of active research. This issue came up in previous discussions, and we actually have an entire separate manuscript devoted to an analysis of the responses to the subjective difficulty question and how subjective difficulty affects the scale factor and the parameters of the utility index of the choice model (Duquette et al. 2009). A draft of this paper is available at http://www.uoregon.edu/~eduquett/research/SCD050109.pdf

In a nutshell, the Duquette et al. paper considers subjective choice difficulty versus objective choice complexity and focuses on proximity of alternatives in utility-space, as opposed to attribute space, in explaining choices. The likelihood function is very balky, since the utility parameters are estimated directly as well as being employed in a construct to measure closeness of alternatives in utility-space, so an alternating maximum likelihood estimator is employed. This other paper falls under the heading of "behavioral economics" and addresses an important issue in the psychology of choice behavior.

As suggested previously, correct signs and significance are reassuring about the structure of the data, but do not rule out the hypothesis that people simplified the task by, for example, recoding

the risk levels to “low,” “medium,” and “high.” A standard scope test would detect such behavior, but correct signs and significance would not.

> *Response*: Again, we believe that our survey design, maximally randomized across individuals, actually produces a standard scope test that vastly exceeds the sort of two-group split-sample scope test that this referee seems to be looking for. Perhaps we misunderstand this point, but we believe this concern is misplaced. See Appendix C (*Details of the Choice Set Design*) for a thorough description of the design of our choice sets.

Given the experimental design used, which includes a constant opt-out reference condition, there are likely to be literally thousands of subject-level internal consistency tests in the data. One such consistency test takes the following form:

Question i: Alternative A_i, Alternative B_i, Neither. Subject chooses A.
Question j: Alternative A_j, Alternative B_j, Neither. Subject chooses B or Neither.

If A_j has the same or better attribute levels than A_i and B_j has the same or worse attribute levels than B_i, then the utility difference between A and the other alternatives is larger in Question j. Choosing B or Neither in Question j is a test failure.

There are several similar tests that occur naturally in the kind of choice designs used in this study that provide direct utility-theoretic measures of data quality that do not rely on indirect checks based on self-reported information.

> *Response*: We had hoped to be able to implement a program that would check for this type of inconsistency (it is one of the options in F. Reed Johnson’s VALIDTST program). The VALIDTST program is configured to work for standard conjoint choice set designs. Unfortunately, it seems we cannot apply this test, or any of the others, with our choice sets because of the way we designed the mix of attribute levels for each alternatives. In brief:
>
> 1. We use too many different levels for our attributes;
> 2. There are no repetitions of the identical choice set for any individual (or likely even across individuals, among the 7520 different choices used in our study);
> 3. Strict dominance of any one alternative over the other in a pair was ruled out in our choice set designs;
> 4. For any individual, the same program alternative was never presented along with different second alternatives (each of the ten programs considered by each individual in their five choice sets addressed a different health threat);
> 5. Utility appears not to be linear and additively separable in each of the attributes in our case. It is not possible to designate individual attributes as conferring unambiguously increasing or decreasing utility, as required for the program. For example, lost life-years are viewed as bad, but for a large enough number of prior sick-years, as in Alzheimer’s Disease, it is possible that more lost life-years would be preferable to fewer.

In our effort to permit valuation of as many different illness types and illness profiles as possible, in no case did we ask the same individual about exactly the same illness profile more than once. Likewise, in order to maximize the preference information revealed through each choice by each individual, we ensured that there was no "strict dominance" between the two program alternatives in any choice set. (By "strict dominance" we mean a randomly generated choice scenario where one program had both a greater risk reduction and a lower cost than the other, since these are the two most-important attributes.) Choice scenarios with strict dominance were rejected, and the randomization was repeated until the two health threats with their risk reductions and program costs would always force the individual to trade off cost against effectiveness.

For each individual, every risk reduction program also addressed a different named health risk (respiratory disease, colon cancer, etc.). Each health risk had a randomly assigned latency, duration, symptoms/treatment, and prognosis (recovery or death, effect on life expectancy). Since all of these illness attributes are randomized, and since there are a considerable number of them, it is hard to see how one would conduct these simple tests.

Ex post, now that we know that there is no problem in identifying strongly statistically significant estimates for the key coefficients, it is clear that we would have been able to afford to repeat alternatives and check for internal choice consistency. This would be easier to do, of course, with fewer attribute levels.

Because of the importance of this issue, we now mention these internal consistency tests in a footnote to the last paragraph of Section I: "A wide variety of non-parametric internal consistency tests can be applied to stated-preference data in some types of applications, but our data involve no repetitions of alternatives within individuals, no instances of strict dominance in any choice set, an attributes for which utility is not necessarily monotonic in the attribute level. The VALIDTST program by F. Reed Johnson permits six types of consistency tests for conjoint choice data. In Appendix B [(*Stated Preference Quality Assurance and Quality Control Checks*)], we describe these tests in greater detail and explain why the VALIDTST program cannot be implemented with our data." Section 11 of Appendix B reviews each of the VALIDTST consistency tests in the context of our dataset.

Certainly, any researcher who contemplates a subsequent study of this nature might wish to seed the scenarios with replicated alternatives for each individual so that internally consistent choices could be tested using VALIDTST. However, it does not seem to be possible with our data. We essentially traded that opportunity for a set of conditions that would maximize our chances, *ex ante*, of finding statistically significant coefficients. Even that effort might have been enhanced had we used a small number of discrete levels of each attribute and pursued "efficient design" strategies (within each of the 138 distinct gender/age groups in our sample)—although most of these rely upon one's priors about the true coefficients, and there are too few previous studies that produce priors for the coefficients featured in our models. Of course, our study produces such priors for use in future studies.

It is possible, given the various efforts the authors made to overcome known problems in stated preference surveys, that most of their subjects who performed pretty well on these internal-validity tests and that including those who didn't would have relatively little impact on their

estimates. In the absence of a quantitative assessment of the utility-theoretic properties of their data, however, it will be difficult to allay concerns about the face-validity of the choice task.

> *Response*: This comment again seems to presume that we can offer only "internal-validity" tests. We believe that this concern is misplaced, since our choice scenarios are randomized across individuals as well as within individuals.

Survey Design

Good-practice reporting for stated-choice studies includes a discussion of experimental-design construction and design properties. Even if nothing particularly sophisticated was done, they probably didn't pay much of a penalty for an inefficient design given the size of their sample. On the other hand, orthogonality may be more of a concern, since, even if they started with an orthogonal array, they removed some implausible profiles from the design. It would be appropriate to reassure readers that the resulting effects on orthogonality were minor.

> *Response*: Thank you. We agree that it is a very good idea to fully describe the construction of our stated choice scenarios. We now include an extensive appendix which details the randomization of choice sets across individuals and across choice sets for each individual: Appendix C (*Details of the Choice Set Design*). We address the question of "practical" attribute orthogonality explicitly in Section 15 of Appendix C. We sense that there may be some difficulty here between the notion of orthogonality in the general sense (meaning simply a lack of correlation, or minimal multicollinearity) and orthogonality in the sense it is used in the experimental design literature (e.g. "Latin squares", etc.)

Health outcomes are labeled as prostate cancer, breast cancer, colon cancer, skin cancer, lung cancer, heart disease, heart attack, stroke, respiratory diseases, diabetes and Alzheimer's disease. (That is eleven, not twelve outcomes. There also was an injury outcome.)

> *Response*: Our count of twelve health outcomes includes the injury outcome, where death can be a consequence as well. Prostate cancer is raised only for males, and breast cancer only for females. Thus each individual, conditional on their gender, could see eleven possible labels for different health threats. Table C1 of Appendix C (*Details of the Choice Set Design*) lists all twelve health threats. (This is our former Table 1.)

The decision to use labels rather than the more commonly applied objective descriptions of disease states deserves some discussion. Some researchers have found that subjects apply a wide range of disease states to the same label, depending largely on their personal experience with the named condition.

> *Response*: Yes we have appreciated this concern for some time. This is why we *also* provided a wide range of additional information that constitutes a more detailed characterization of the illness profile associated with the label. See Appendix C (*Details of the Choice Set Design*) for the full list of attributes that we vary with each label-

specific illness profile. Also see Appendix A (*Survey Design & Development*) for a more general description of the survey choice sets.

We used labels for each health threat because we wish to know whether the label itself conveys information that is distinct from the actual illness profile. This type of information is needed to permit us to determine whether there is a "premium" for different types of illness. It has been a long-standing policy question, for example, whether there should be a "cancer premium" in the *VSL*. We have another distinct separate paper (Johnson et al. 2009) that adds indicators for the randomly assigned illness labels as shifters on the various basic utility parameters in our models. This paper can be accessed at
http://www.uoregon.edu/~cameron/vita/Cameron_DeShazo_Johnson_0619091.pdf

We find that merely the label associated with the illness can produce a "lump" of additional utility associated with a reduction in the risk of that illness. Likewise, the marginal utilities of avoided sick-years and avoided lost life-years are allowed to shift with these illness labels, producing a number of statistically significant effects. We use these disease-label models to produce estimates of *VSL*-like quantities that are illness-specific. One interesting finding is that non-smokers are willing to pay very little to reduce their risks of respiratory disease and lung cancer, but smokers have much higher than average *WTP* to reduce these risks.

There is a similar problem with recovered/remission status. The authors conclude that "most people seem to associate recovered/remission status relative to any of these major health threats as involving considerable limitations." In other words, this outcome is not under experimental control, since quality of life is not explicitly specified. Thus people will imagine different outcomes, which induces measurement error in the estimated parameter.

Response: Unlike the illness labels, we struggled with how to handle this attribute. In practice, people will make decisions about risk reduction options for serious illnesses in real choice contexts where they are *also* not entirely clear on their prospects for recovery or their quality of life should they suffer such an illness and recover.

When we were designing our survey, we contemplated telling respondents to assume that "recovery" from one of these major and life-threatening illnesses should be interpreted as restoring them completely to their current health state, but this was clearly implausible. For example, the life of a cancer survivor will be different from that of a person who has never had cancer. If people did assume that recovery was complete, we would expect to find that the utility *differential* for being in "recovered/remission status," relative to their present health state, would be zero. Whenever we detect a statistically significant coefficient for this differential, though, it is negative. This appears to support our conjecture.

We certainly agree that bringing this variable more under experimental control would be desirable. In our separate "disease label" paper, we introduce several variables that capture individual heterogeneity in risk perceptions: controllability of the disease in question, and the individuals subjective risk of experiencing a future bout of the disease in question, along with their faith that medical care they receive would be timely and effective. We have focused on shifts in baseline utility levels, and in the marginal (dis)utility of a sick-year or a lost life-year. Perhaps we should delve more deeply into the

> marginal (dis)utility of a recovered/remission year. We do have information about whether the individual has friends or family who have suffered from each named disease. This familiarity may improve the individual's assessment of their likely quality-of-life in the recovered/remission state.

The risk-reduction commodity is described as "purchase of new early diagnostic programs that would be coming on the market … . These programs are described as involving annual diagnostic testing and, if needed, associated drug therapies and recommended life-style changes." The details of these programs were left up to the subjects' imagination. The authors concede that the timing, invasiveness, frequency of the tests, possible side effects of drug therapies, and required life-style changes could very well involve time, inconvenience, and money over and above the nominal program price attached to the scenarios. The authors felt they could not accommodate these concerns within the practical constraints of the survey and must "assume that the stated cost of achieving the advertised risk reduction subsumes all market and non-market opportunity costs perceived by the respondent." Even given survey constraints, it would have been possible to include a statement explicitly acknowledging such indirect costs and asking subjects to think about the scenario price as including the total personal cost of the program to them.

> *Response*: We concede that, given the opportunity for a "do-over," we might have approached this portion of the survey differently. On **Form 17**, we state specifically that the risk reduction programs in question would *not* involve "uncomfortable procedures." We do state that "Your participation in a program would cost you money." These programs would not be covered by the respondent's current health insurance. "These higher costs might take the form of a co-payment when you visit your doctor or higher monthly health insurance costs." "To make it easier to compare, we present all costs as monthly costs, and also as annual costs. You would need to pay for, and participate in, a program for the next __ years to get its benefits." (The precise number of years corresponded to that individual's current age and nominal gender-specific life expectancy.) We did not explicitly limit the cost of the *program* to simply the cost of the *test*. Instead, we were careful to refer to the "cost of the program" (where the programs are described on **Form 17** as involving prescribed "medication and life-style changes that reduce your risk of getting the illness").
>
> Earlier in the survey, however, on **Form 7**, we specifically asked respondents to consider the difficulty of making life-style changes. We asked them: "Changing your lifestyle or habits can be difficult because it requires time, money, and effort. How difficult would it be for you to do the following things?" The listed options included: drink less alcohol, quit smoking, eat a healthier diet, see a doctor more regularly, exercise more, lose weight, use a seatbelt more. We went through one phase of survey development with language in the instrument where we tried to explain the idea of the monetized disutility of the tests themselves, and opportunity costs and the full cost of time. However, without getting into discussions of the value of travel time to the doctor's office and the pharmacy, and the prospective disutility of a new exercise regimen or dietary restrictions, there seemed to be no happy medium, so we opted for a minimalist approach. Perhaps there would have been a better option, but we could not see it at the time. To meet the length/duration restrictions under our contract with Knowledge Networks, of course, it was

necessary to prune many things out of the survey that we were keenly interested to include. This is a frequent problem with survey research in general.

In response to this concern, however, we have investigated additional models where we allow the estimated marginal utility of net income to depend on the respondent's answers to our questions about the difficulty of accomplishing lifestyle changes. We take advantage of the wording on **Form 7** in the question: "Changing your lifestyle or habits can be difficult because it requires time, money, and effort. How difficult would it be for you to do the following things?" A slight complication is that respondents were only asked about each of these things if they responded on **Form 6** that there was still at least some room for them to reduce their health risks by improving their lifestyle or habits in these ways. We assume that if the individual reports no room to improve along any particular dimension, then it would be very hard at the margin for them to improve any further on this dimension. (Cleaning up a few of one's bad habits may be relatively easy, but getting rid of all of them might be tough.)

However, if there is still room to improve on one or more dimensions, and respondents report that it would be easy or difficult for them to do so, this is the notion we wish to capture. We construct a crude variable to measure "ease of improving health habits." For each type of the seven health habits identified on **Form 6** and **Form 7**, we build two variables. One is prefixed by "improve_" and measures "opportunity for improvement" with ratings that vary from 0 = "no opportunity for improvement" to 4 = "much room to improve." The second variable is prefixed by "easy_" and measures the ease with which these available improvements in health habits could be accomplished. For this variable, we have inverted the question about how *difficult* it would be to make improvements. For our "easy_" variables, the ratings are coded as 0 = "hard to improve" to 4 = "easy to improve."

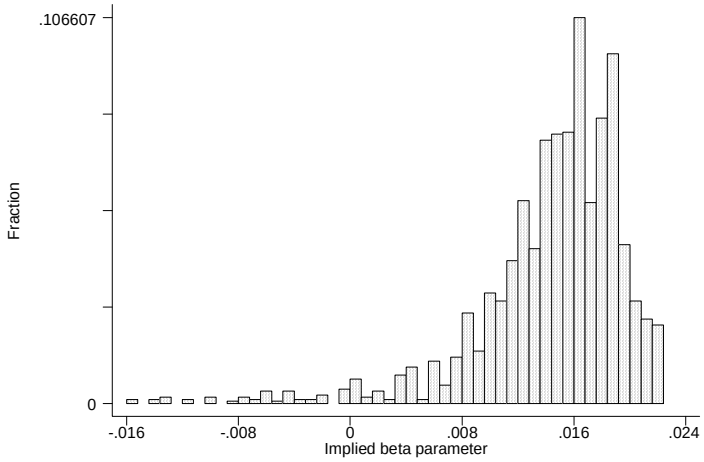
For each of the seven health habits, we construct an interaction between the "improve_" and "easy_" variables. This interaction term is zero if the individual has no opportunity to improve *or* if they do, but it would be very hard for them to do so. This interaction term takes on a larger value (to a maximum value of 16) if there is lots of room for the individual to improve their health habits and they believe it would be easy to do so. Acknowledging the degree of approximation involved in the use of ratings, and the different metrics across the different questions, we then forge ahead and add these interacted ratings across all seven types of health habits to generate a variable that may serve as a proxy for the likely psychic or non-pecuniary costs to the individual if they need to make "lifestyle changes" in addition to paying for the annual pin-prick blood test in the choice scenarios.

The maximum value for our constructed indicator is 16x7 = 112. It measures "ease of making lifestyle changes." We desire a variable that will be larger if the implicit costs to the individual of making these changes is larger, so we subtract our indicator from 112 to convert it into an indicator called $hard_i$ which proxies for the "difficulty of making lifestyle changes." As a further complication, however, not all respondents answered all of the questions on **Form 6** and **Form 7**, so we create an indicator for whether information was missing. 1,724 of our 1,801 respondents provided sufficient information to build this variable. We thus use a second indicator variable to control for data availability.

Now we simplify the intuition by supposing that the indirect utility difference that drives program choices is linear in net income and we don't need to worry about the pattern of net

income across the uncertain prospects of getting sick or remaining healthy. In that simple case, $\beta(Y-c)-\beta(Y)=\beta(-c)$. Suppose costs are perceived as systematically higher than what is stated in the choice scenario, say $c\theta$, where $\theta>1$. If respondents are reacting to this larger cost, but we control only for c, then we will actually be estimating $(\beta\theta)(-c)$, and the apparent "marginal utility of net income" coefficient will be too large. This coefficient forms the denominator of the WTP function, so a too-large value will lead to a WTP estimate that is too small. People who look like they are unwilling to pay the amount stated in the choice scenario are actually unwilling to pay the larger amount, $c\theta$. Failure to accommodate these other implicit costs will lead to underestimates of WTP.

We incorporate our new variable, $hard_i$, along with the indicator for its availability, into our model by allowing these two variables to shift the β coefficient. The slope coefficient on the interaction with the indicator variable is insignificant, but the slope coefficient on the interaction with $hard_i$ is positive and strongly significant.



There is a range of perceived difficulties of making life-style changes among our respondents. The values of the $hard_i$ variable range from 0 through 112 , with a median of 92 and an inter-quartile range of 82 through 99. For people who perceive life-style changes as relatively more difficult (i.e. those who may consider other implicit costs associate with each risk reduction program), the marginal utility of income is estimated to be higher, which would imply a lower WTP for the risk reduction programs in the choice scenarios. For people who perceive life-style changes as relatively easier, the marginal utility of income is estimated to be lower, which would imply a higher WTP for the risk-reduction programs in the choice scenarios. Appendix E (*Model, Estimation and Alternative Analyses*) includes an alternative definition of the variable $hard_i$.

The relevant question, now, is "what would people have been willing to pay had they believed that the quoted cost on the survey was the full cost of the program—i.e. that there were no additional costs associated with the difficulty of complying with the lifestyle changes that might be required?" It might be tempting to simulate the value of the marginal utility of income parameter for the case where everyone believes that it is trivially easy to implement life-style changes. This would correspond to the counterfactual where nobody perceives any implicit costs of this variety in addition to the cost of having the test.

We had intended to do this sort of thing in our analysis, which was why we collected the information on **Form 6** and **Form 7**. However, we did not anticipate that respondents might view "lifestyle changes" in two separate ways. We expected that people would view them as necessary *complements* to the health testing programs described in the choice scenarios. This is the implicit assumption behind the concern that respondents will impute other costs to each program besides just the cost stated in the survey question.

However, it may actually be the case that respondents view the testing programs in the survey as *substitutes* for the lifestyle changes that they know they should really be trying to make. If they perceive that participation in these testing programs will allow them the luxury to continue with their current poor health habits but still lower their health risks, they may actually express greater demand—because the perceived benefits are greater than just the reduction of health risks.

This makes things considerably more complicated. If we were to simulate a situation where everyone found it perfectly easy to implement any life-style changes that would be required along with the testing program, the marginal utility parameter for income would be vastly smaller, causing the inferred WTP for these programs to *increase*. But here's the catch: if life-style changes were easy, the "price of a substitute" for the testing program would also be dramatically smaller, which would *decrease* the demand for the testing programs. People could simply change their health habits and they would have no need for the testing program. Thus it seems highly inappropriate to consider any adjustments to the stated cost of the program without making corresponding adjustment to the price of substitutes. Clearly, more research is needed, and it should focus on this "complements versus substitutes" distinction.

Incidentally, we do have some evidence, in other work with these data, for the "substitutes" possibility. In our research concerning the disease labels, non-smokers are willing to pay very little for tests to reduce their risk of lung cancer or respiratory disease, whereas smokers are willing to pay amounts for these two illnesses that substantially exceed the WTP amounts measured for all other illnesses for the general population. In this case, the substitution effect appears to dominate very strongly.

We have incorporated the text of this response, with a few embellishments, into Appendix E (*Model, Estimation and Alternative Analyses*).

The authors inform subjects that "they have an existing risk of suffering from the illness or injury in question." Is there any evidence that subjects accepted risk endowments that may vary considerably from their own risk perceptions?

Response: Yes, we do have some evidence and in some cases respondents appear to subjectively "adjust" some the information we gave them. Fortunately, we ask respondents about some (although not all) of these adjustments. Respondents were asked, on a program/disease by program/disease basis, about whether they expected that this specific program would ever benefit them, and if so, at what point in the future. The "scenario adjustment" paper (Johnson et al. 2009) is devoted to analysis of these questions and how respondent heterogeneity along these dimensions affects implied *WTP* for risk reductions. This paper can be accessed via http://www.uoregon.edu/~cameron/vita/

A different paper (DeShazo and Cameron 2005) explores systematic differences in *WTP* by same-illness current morbidity and other-illness (comorbidity) status. This paper can be accessed at http://www.uoregon.edu/~cameron/vita/comorbidity_121105.pdf . Risk perceptions do indeed have statistically significant effects on the marginal (dis)utilities of adverse health states. As mentioned above, we rely on the randomization of illness names and attributes to minimize omitted variables bias. Without controls for subjective illness risks, of

course, we estimate sample *average* marginal utility parameters that reflect the distribution of preferences in the sample (and hence in the general population, since our sample is representative). If we do control for subjective risks, we can quantify a lot of systematic variation in marginal utilities. Of course, these subjective risks are not generally observable. Our survey data are unusual in this respect.

The model is laid out for the two-alternative case, with the comment that "the three-alternative case is completely analogous." In fact, adding a third, fixed opt-out or no-purchase alternative introduces several complications of empirical concern.

Response: The initial two-alternative case used to develop the model is a choice between the respondent's (stated) current risk of experiencing a given illness profile at zero cost versus a reduced risk of experiencing that illness profile at positive cost. The "third alternative" that we add is *not* a "no purchase" alternative (as suggested in this comment), but rather *another intervention program* that addresses a different illness with a different baseline risk (which provides a different risk reduction at a different cost). Health risks from these illnesses are present at different levels under all three alternatives. For two of the alternatives, one particular health risk is reduced at a cost.

At the urging of the editor, we now include Appendix E (*Model, Estimation and Alternative Approaches*). This appendix includes a more detailed set of steps that explains how we arrive at our estimating specification and the various formulas summarized in the main paper. It also includes both the biostatistical and econometric views of fixed effects conditional logit models. Appendix E also includes the full three-alternative, five-choice likelihood function, but outlines how it is actually estimated as a conditional likelihood, to permit any unobserved interpersonal heterogeneity (in the form of fixed effects) to be netted out.

We also undertake a standard Hausman test for whether the changes in the remaining parameters are sufficient to suggest that the fixed effects model is necessary. These parameter differences do not indicate a great need for this generalization (although the constructed covariance matrix for the difference in the estimates is unfortunately not positive definite). This is not too surprising, since the attribute levels are randomized. It is therefore unlikely that these illness profile attributes will be correlated with many unobserved individual characteristics. We control explicitly for age, which is the main conditioning variable for admissible illness profiles, so differences in respondent ages are not "unobserved" heterogeneity.

First, the third alternative will violate the IIA assumption of conditional logit.

Response: *A priori*, we would not expect a second risk-reduction program alterative to be an "irrelevant alternative" in the sense that choice researchers worry about the Independence from Irrelevant Alternatives problem in multiple choice contexts. An irrelevant alternative typically differs only along dimensions for which marginal utilities are zero (i.e. the classic red bus, blue bus problem). The second program option offered in every choice set differs on all attributes from the first program option, and even the basic models in Table 2 reveal that all of the key attributes of our programs matter very significantly to people's choices. The t-test statistics on the slope coefficients in Table E1

in Appendix E (*Model, Estimation and Alternative Approaches*) are large by anyone's standards. These attributes have a very significant effect on people's choices. This would seem to indicate that a second program alternative is not "irrelevant."

Second, the third alternative compromises the orthogonality of the two-alternative experimental design.

Response: This concern may stem from the lack of a clear enough description, in the earlier version of the paper, of the nature of the randomizations used to generate our 15,040 different combinations of illness profiles and risk-reduction programs. Each respondent needed a set of illness profiles tailored to their age and gender, so we gave up on the prospect of building a conventional "experimental design" with a finite number of attribute mixes. Table C2 in Appendix C (*Details of the Choice Set Design*) now contains the entire joint distribution of ages and genders in our estimating sample. We would have needed 138 different sets of "experimental designs" to accommodate both age and gender, and some group sizes are much too small. Genuine randomness, subject to a few plausibility constraints, seemed the most straightforward way to proceed.

In our two-illness-profile-plus-status-quo framework, each individual faces a whole spectrum of health risks under the status quo—for each relevant health risk of the twelve types of health risks we highlight in our study, as well as many, many others. Each choice scenario focuses on two specific health risks out of the constellation of risks faced by each individual. In the single example of a choice set featured in the appendix to the main paper, the baseline risk for the "Heart Disease" threat is given as 40 in 1,000 and the baseline risk for "Colon Cancer" is given as 4 in 1,000. These are the risks that the individual is quoted for these two health threats if the "Neither Program" option is chosen. If the individual chooses the "Heart Disease" program, however, two things change: the risk of heart disease drops from 40 in 1,000 to 38 in 1,000 and the individual's net income falls by $180 per year ($15 per month). Similarly, if the individual chooses the "Colon Cancer" program, two things change: their risk of colon cancer drops from 4 in 1,000 to 2 in 1,000 and their net income falls by $48 per year ($4 per month). The two programs are thus not drastically different from the status quo. It is not the case than any health risk is changed dramatically, and certainly not to zero. All of the alternatives involve slight variations in health risks and modest variations in net income.

The concern raised in this point may be an artifact of the referee's mistaken impression that every respondent saw the identical set of five choice scenarios, and that these scenarios varied only across choice sets, within identical survey instruments. As detailed above, this is not the case. Every single illness profile was randomly generated. Individual disease profiles were screened for implausibility and re-randomized as necessary. The construction of choice sets from pairs of programs involved screening of each pair of illness profiles to preclude strict dominance in terms of risk reductions and costs. Across the ten illness profiles shown to each respondent in their five choice sets, no illness label was duplicated and pairings were random (subject to a constraint that "Traffic Accident" was not among the first two randomly labeled illness profiles which were also used in the tutorial portion of the survey). In no case did an individual view an illness that would affect them sooner than two years from now (i.e. all illness profiles

were based on the respondent's own current age and nominal life expectancy and were represented as future illnesses). The age at which they would get sick was randomized for each illness profile and the nature and duration of symptoms and treatments were also randomized. Age at recovery (if any) was defined by the randomized duration of sick-time. Age at death was also randomized, defining the period in recovered/remission status and the number of lost life-years relative to nominal life expectancy.

Third, differences in the scale of the A/B alternatives and the alternative-specific, opt-out constant affect direct utility comparisons.

Response: As indicated in the second paragraph of our previous response, above, the differences between the A/B alternatives and the status quo alternative are not particularly drastic. All three alternatives involve a wide variety of existing health risks. Alternatives A and B involve slight reductions in two of these, each at the cost of a modest reduction in annual net income.

With three alternatives, of course, it is possible to explore further *ad hoc* generalizations of a basic conditional logit specification. The models we report in the main paper involve no alternative-specific constants. Our theoretical choice model assumes that utility depends simply upon health status in each future year and on net income, and respondents' choices depend on their discounted expected streams of utility under each alternative.

Since we do not use an "alternative-specific, opt-out constant," it seems that the referee may be worrying about something that is not relevant to our specification. Every alternative shares the same set of utility parameters, so there appears to be no differences in scale, *a priori*, to contend with.

Of course, one can certainly introduce an *ad hoc* alternative-specific dummy variable that allows for an extra increment of unspecified utility (positive or negative) for the "Neither Program" very "Any Program" alternatives. We include the results for such a model in Section 6 of Appendix E (*Model, Estimation and Alternative Analyses*). However, it is hard to say in advance how this increment would be explained. It can be interpreted as the *net* effect of host of potential biases, including status quo bias, payment vehicle rejection, yea-saying or nay-saying. The utility-maximization story does not provide a specific rationale for a lump sum of autonomous utility associated with the status quo or "any program" in people's tradeoffs between net income and health risks, although all of these potential behavioral nuances often mean that either a positive or a negative and significant coefficient can be estimated on a status quo or "any program" dummy variable. We could digress to include this model in the paper, but given the length constraints, it seems preferable to cover this issue in Appendix E.

Estimation

There is virtually no conceptual model behind the analysis other than the standard expected utility hypothesis and implicit utility axioms underlying welfare theory. There is an extensive literature that stated risk preferences violate expected utility and that various forms of rank dependent utility describe choice data better.

Response: Given the complexity of our survey instrument and estimating specification, it seems prudent to start with a simple model like expected utility for the analysis of choices such as those we consider here. A considerably longer and more complicated paper could entertain a selection of non-expected utility specifications, in addition to this starting point, and use non-nested tests to ascertain which type of specification best explains the choices made by our respondents. It seems like this would be a very different and perhaps subsequent paper, referring to this one. Such a paper would also need to cover the differences in welfare calculations in non-expected utility contexts, for example.

At the beginning of Section II.A., we now include a footnote that reads: "The literature has identified a number of anomalies that cannot be explained by conventional expected utility models and exponential discounting formulas, but these specifications can serve as a useful starting point for our analysis. Subsequent research with these data may explore non-expected utility models and different types of discounting."

While the authors have employed advanced econometric methods to estimate a choice model using a flexible utility specification, I also would like to see the results of a mixed-logit categorical model that imposes no functional-form restrictions at all.

Response: A mixed-logit model has been estimated, corresponding to Model 2 (formerly Model 7), where the six baseline parameters (β_0 on the net income term, α_{10} on the sick-years term, α_{20} on the recovered/remission years term, α_{30} on the lost life-years term, α_{40} on the squared term in lost life-years, and α_{50} on the interaction term between sick-years and lost life-years) are all allowed to be random while the remaining parameters are fixed. Note that the distribution of the age variable across our sample can be approximated with a normal distribution. When these six parameters are allowed to be normally distributed, it is not surprising that the interaction terms involving the age variable become statistically insignificant. In parsimonious models that drop the insignificant age shifters, the first four mixed-logit coefficients display robustly statistically significant variances across the population: β_0 on the net income term, α_{10} on the sick-years term, α_{20} on the recovered/remission years term, and α_{30} on the lost life-years term.

In our analysis, however, we prefer *explained* parameter heterogeneity to the *unobserved* parameter heterogeneity captured in random coefficients models. The respondent's current age contributes significantly to explaining heterogeneity in choices, and *WTP* for risk reductions as a function of age is an important policy question, so we prefer a model with age as an explicit source of heterogeneity.

We now discuss mixed logit models in Section 5 of Appendix E (*Model, Estimation and Alternative Analyses*) including a histogram for the age distribution in the sample.

Are the attribute-level parameters naturally ordered? Are adjacent parameters significantly different from each other? Are utility differences commensurate with numerical differences in the risk and price levels?

Response: Other than the disease-label attribute, most of the attributes are naturally ordered because they are cardinal variables. As clarified above, unlike many conjoint choice experiments, our design involves random draws from *continuous* distributions for each attribute (other than the twelve disease labels). In simpler designs with just three or four levels per attribute, it is highly appropriate to estimate utility differentials for each level relative to a base level. We have many levels per attribute, as described in detail in Appendix C (*Details of the Choice Set Design*). In our case, we would have to lump these continuous variables into bins to avoid functional-form restrictions.

The translog-type form is offered as a second-order local approximation to an arbitrary functional form, which is generally assumed to be fairly accommodating, especially in the presence of interaction terms. The linear-in-attributes *ad hoc* and structural forms in Table E1 in Appendix E (*Model, Estimation and Alternative Analyses*) are provided to show that the most important variables are strongly significant even without the greater generality provided by the translog-type form.

The authors use a square-root transform of income using the following logic. A line-search across Box-Cox transformation parameters maximized the log likelihood at 0.42 and there is a negligible difference between the maximized log-likelihood values for lambda values of 0.42 and 0.50. Unfortunately, the square root is not the same as Box-Cox lambdas of either 0.42 or 0.50. Assuming income is measured in thousands, the following table suggests that the closest approximation to square root is a Box-Cox lambda of about 0.28.

Y	sqrt	λ=0.5	λ=0.42	λ=0.28
10	3.2	4.3	3.9	3.2
20	4.5	6.9	6.0	4.7
30	5.5	9.0	7.6	5.7
40	6.3	10.6	8.8	6.5
50	7.1	12.1	9.9	7.1
60	7.7	13.5	10.9	7.7
70	8.4	14.7	11.8	8.2
80	8.9	15.9	12.6	8.6
90	9.5	17.0	13.4	9.0
100	10.0	18.0	14.1	9.4

Response: We appreciate the effort gone to by this referee to check this point numerically. However, the assumption that we measure income in thousands is incorrect. Income is measured in dollars, which means that the approximation is actually very close. Since this concern has been raised, however, we have undertaken to re-estimate the original Models 3 through 7 and redo all of our tables, simulations and figures using the likelihood-maximizing Box-Cox lambda parameter of 0.42. Our new Table 1 now shows that the maximized value of the log-likelihood is minimally higher (by 0.36) with a Box-Cox parameter of 0.42 (rather than the more-restrictive square root transformation). Parameter differences are very modest as well (typically by no more than one digit in the second or third significant figure). There are certainly no qualitative differences in our estimates.

Less serious distortions are related to shifting the data for each present discounted value term by one to avoiding taking the log of zero. A rule of thumb for log transforms sometimes used in regression models is to recode zeros to half the distance between zero and the lowest positive value in the data. Since potential distortions occur primarily at small values, this decision may not be as consequential as the square-root transform.

> *Response*: Yes, the referee's concern would be warranted if we were shifting *only* the zeros in these data. However, we shift *every* value of the variable by one unit before taking logarithms, as indicated by our use of $\log(pdvi_i^j + 1)$ in the utility-difference formulas. Our rationale was to have the *transformed* variable equal zero whenever the *raw* variable equals zero, since there should be no utility difference relative to the status quo when there is no illness (i.e. if sick-years are zero, recovered-years are zero, and lost life-years are zero, the person remains on the trajectory they would have enjoyed without this particular illness). It is of course true that potential distortions under log transformations occur primarily at small values, since the logarithmic transformation converts small fractional values into large negative numbers. After every observation has been shifted by one unit, however, there are no values less than one in the data. This strategy to avoid the "log of zero" problem appears to be quite common. (For example, we notice that it is used in the recent Dobkin and Nicosia paper concerning the War on Drugs, *AER*, March 2009).
>
> We now include in Appendix E (Model, Estimation and Alternative Analyses) a discussion of our rationale for using the logarithmic transformation that we choose for the discounted health state duration variables. We impose the same transformation for all three discounted health durations, and conduct a line search across possible values of a Box-Cox transformation parameter these variables. Section 6 of Appendix E includes Figure E2, which shows that the log likelihood improves by five points as the shared transformation parameter is reduced from one (corresponding to a linear transformation) to zero (corresponding to the log transformation used in the paper for all three terms). When the parameter is reduced below zero, the log-likelihood continues to improve a little, but much more slowly. There is only a one-point improvement between zero and -1. Thus we decide to stick with the familiar shifted logarithmic transformation for these variables.

Results

The authors deserve credit for trying to rid us of the widely misunderstood and misused term "value of a statistical life." Nevertheless, after arguing against the *VSL* terminology they embrace the term "value of a statistical illness profile". Moreover, after criticizing the implicit linearization and inappropriate extrapolation from data on the value of small risk reductions to one in *VSL* calculations, they calculate *WTP* for a micromort. Since their subjects evaluated changes much larger than 1 in a million, the same criticism about inappropriate extrapolation could apply.

> *Response*: This point is well taken. Our concern in extrapolating *WTP* to a risk change of 1.0 stems from the public's confusion that this yields the "worth of a human being." In our study, the baseline lifetime risks vary between 0.004 and 0.04. The risk changes in

question range from 0 (for the "Neither program" alternative) to between -0.001 and -0.006 for the different risk-reduction programs. Lifetime risk *changes* between 0 and 0.006 are, therefore, represented in our sample, but we do not have non-zero risk changes smaller than 0.001 in our data.

Thus, we cannot say with certainty that there is no threshold below which people are willing to pay zero to reduce a risk, or that people are not willing to pay positive amounts to reduce even infinitesimally small risks. If *WTP* is not proportional to risk changes, even at low levels, then any arbitrary benchmark for normalization is inappropriate. (It will not be surprising that we have another paper in the works that examines the proportionality assumption and baseline risk effects using these data.)

In our revisions to the paper, we have eliminated the *VSIP* term. We now use the term "*WTP* measure". On the matter of micromorts, it might be preferable to use a millimort (a one-in-one-thousand mortality risk change), but for our study, risks are larger because they are quoted to respondents as *lifetime* risks, rather than annual risks. In the more conventional wage-risk *VSL* context, risks are typically smaller *annual* mortality risks in specific jobs (and the wage differentials necessary to staff these jobs are annual wages). When we introduce micromorts, we include a footnote describing the numbers of micromorts at stake in familiar wage-risk contexts.

Contrary to previous studies, the authors obtain an increase, rather than a decrease, in *WTP* among seniors. Some discussion of the possible empirical or theoretical significance of this result would be useful.

Response: We obtain the usual result, shared with earlier studies, that seniors are less willing to pay than middle-aged adults for reductions in the risk of "sudden death in the current period" that is implicit in most *VSL* estimates. However, Figure 3 in the paper (the one which *does* show WTP increasing with age) is for an illness profile that earlier models have not been able to address. This illness profile involves getting sick one year before the end of the individual's nominal life expectancy, spending a half a year sick, then dying six months earlier than would otherwise have been the case.

The energy and creativity devoted to this study is impressive. The authors' arguments for allowing for intertemporal effects and substitution opportunities are persuasive. Therefore, it is disappointing that their model "produces *VSL*-type estimates which are squarely in the range produced by other studies." This surprising result undermines their otherwise persuasive case that it is important to do things correctly. Apparently it really isn't so important, after all.

Response: This perception has greatly helped us in rewriting the paper to emphasize that our data and our model allow us to do far more than simply confirm existing VSL estimates.

First, the value of our approach is not that it provides merely one more estimate of the *VSL*, but that it brings better and more information to bear on estimating the entire schedule of marginal *WTP* for risk reductions across an individual's future life years and sick years. Thus, the value added of our approach is increased validity (better definition of the object of choice), more information (individual preferences determine the value of future risk reductions, instead of just a researcher making extrapolations), and its

comprehensiveness in terms of valuing different types of illness profiles (and thus providing valuation information on policies with different latencies for both morbidity and mortality).

Second, we view the closeness of the benchmarking case as a *strength* of our model. It should be very reassuring to our readers that a special case of our model—for an illness profile that consists of sudden death in the current period—produces estimates that correspond well to the population averages found in one-size-fits-all *VSL* calculations. Fortunately, our goal is not merely to produce "better" estimates of the *VSL*. Instead, our goal is to provide a way to value the benefits of reducing a broad range of health risks, not just sudden death in the current period. That we can easily replicate the magnitudes of *VSL*s in our *VSL*-type special case lends greater credibility to our estimates when our model is specialized to illness profiles *other than* sudden death in the current period—estimates that no earlier study can provide. Tables 2, 3, and 4, along with Figures 2 and 3 are provided to illustrate the much greater generality of our models. The "sudden death" scenario is just one possible illness profile among many (i.e. it is the case where the rest of the individual's nominal lifespan is to be experienced entirely as lost life-years). The rows of Tables 2 and 3 give estimates for four other types of illness profiles, involving different durations of morbidity and different prognoses (i.e. fatal and non-fatal, but serious, illnesses). Table 4 shows the implications of increased illness latencies, for a 35-year old, and then for a 65-year old. Earlier models cannot generate estimates like this.

Other

The authors describe their method variously as "stated-preference experiment," "stated preference survey", and "conjoint choice experiment" at different points in the manuscript. Terminology is not standardized in this area, but my personal preference is to say "stated choice survey, also known as conjoint analysis or discrete-choice experiment."

Response: We are happy to use this terminology as well. We have implemented this in our new introduction. Stated preference researchers in different subdisciplines of economics (transportation, marketing, environmental and health) have evolved different terminologies, so it is important to remind readers that they all refer to the same thing. We now refer to the survey explicitly as a "stated choice survey" and use "stated-preference" as a more general adjective for the type of information derived from such survey.

The authors are a bit too eager to sell their study with terms like "we worked hard", "we carefully explained", "careful survey design", "careful analysis", "this information ensures", their model, "seamlessly" accounts for morbidity and mortality, etc. It would be better to describe their efforts and let readers judge how careful, successful, and seamless those efforts are.

Response: Yes, we would be happy to remove these adjectives and adverbs from the paper. They have crept in merely in response to skepticism expressed by other readers (of this and other papers related to our larger suite of studies) concerning the rigor of our

work. Various papers from our larger project have collectively been through the process of presentation and discussion at several years of conferences, workshops, and seminars. It has always been difficult to describe in sufficient detail, in any one paper, all of the intricacies of survey design, sampling, and administration as well as the substantive points. Much of the language that irritates this referee evolved in response to the comments of readers who expressed skepticism about stated preference methods in general and (if we may be permitted some hyperbole) questioned whether we'd dashed off the survey instrument over a weekend, fielded it to a convenience sample of college students in a few of our classes during the week, and then written it up over the next weekend.

The model notation is rather tough sledding. While this may be inevitable, considering the complexity of the analysis, one wonders if the rather arcane notation is really unavoidable. Given the length of the manuscript, it might be advisable to write out a more simplified model and indicate how it was generalized empirically.

Response: We are not sure how to proceed in response to this advice. We have already pared down the model so that we develop it initially in the context of just two alternatives and for indirect utility that is simply a function of net income and a set of four mutually exclusive and exhaustive future health states. We make liberal uses of abbreviations for complicated terms that tend to appear in multiple places in the equations in the paper. Our new sub-section headings may help readers follow the logic a little more easily.

We use simplifications to suppress the necessary complexity in the constructed variables for the estimating specification that involve the net income terms under each alternative (this complexity stems from the different future profiles of income and program costs when the individual is either sick or dead, with a given probability). Those terms are daunting when expanded, but we have spared the reader that notational complexity in equations (3) and (4). We develop the annual option price in equation (5) and the present discounted value of its expectation (6), given that future illness is only probabilistic. As we develop the model, we use a generic form for the effect of net income on utility, $f\left(Y_i\right)$, and its inverse. To illustrate our closest approach to a conventional *VSL*, we simplify to a very easy case where $f\left(Y_i\right) = \beta_0 Y_i$. This allows us to make our basic point, although our estimating specification uses $f\left(Y_i\right) = \beta_0\left(\left(Y_i^{\lambda} - 1\right)/\lambda\right)$ where a value of $\lambda = 0.42$ is suggested by the data.

To assist the referees and anyone who might be sufficiently interested in the details to wish to re-derive our formulas, we now provide far more details concerning the derivation of our estimating specification in Section 2 of Appendix E (*Model, Estimation and Alternative Analyses*).

I suggest deleting the material on sample-selection bias. The authors can cite their separate publication on this analysis.

Response: The sample selection effect is statistically significant, so it would be inappropriate to deny the data and assume it is zero. However, we are happy to leave most of the discussion of this effect out of the main paper. We have incorporated the

relevant material from our separate paper in our new Appendix D (*The Knowledge Networks Panel and Sample Selection Corrections*). The separate paper is not entirely adequate because it concerns a modified selection equation and a different outcome variable. The selection equations used to build the correction variable used in this paper are more general.

We especially want to thank this third referee for the effort and time they devoted to improving both our paper and our broader study, since the information we have added in response to these questions, in the new appendices, will support most of the other papers from this project that we also have in the works. This referee's input has been invaluable because of their specialized knowledge of the challenges that practitioners of stated choice experiments face, and the standards to which good studies should be held. We've done our best to meet these high standards.

7.2.3 Third round of reviews from the *AER*, with what would have been our replies

As mentioned above, the *AER* editor went to a fourth reviewer (as a replacement tie-breaker) when Referee III abdicated from the process. Below are the comments of referees IV and V (where referee V is a colleague he recruited to assist with the review). Based on this review, the editor decided to reject our paper without inviting a response from us. Nevertheless, we provide here the responses that we would have provided, and note that we have further revised the manuscript to prevent subsequent readers from developing the same sorts of mistaken impressions as these two.

7.2.3.1 Combined comments of AER referees IV and V

The objective of the paper and the research is interesting and important. Health economics definitely needs to reform and replace the idea of QALYs, and this could be a major step in the right direction. In particular, it provides solid theoretical thinking about a new way to approach the problem, but unfortunately, there are issues with the implementation. That said, I think that this work should be encouraged, especially in light of the widespread use to which QALYs are put, and the likely adverse impacts of using them in many cases, especially when the underlying "theory" is at best weak and at worst ad hoc.

RESPONSE: We are grateful for the recognition that it is time for a new utility-based approach to the task of valuing health risk reductions. We concur that QALYs, while solidly entrenched in the literature and in medical cost-effectiveness practice, leave a lot to be desired. We agree that our approach represents a solid and logical theoretical approach to this problem, but we argue below that most of the "issues with implementation" that these referees raise do not seem to be particularly relevant in this application.

Unfamiliar constraints, for realistic choices: In particular, these referees wish we had designed our choice sets differently. However, in our responses below, we must disagree that there would have been any other way to design choice sets that would preserve the key innovations in our study. We suspect that these referees merely overlooked that it is

necessary to index each choice scenario to the individual's particular age and gender. We explain below why this is a crucial and necessary constraint (albeit one that is unfamiliar to marketing researchers), and why any model that focuses on prospective illness *profiles* must accommodate this constraint. This is a requirement for taking "a major step in the right direction," but it simultaneously precludes the design features that these referees wish we had used.

Neither of us "buy" the arguments about the way the discrete choice experiment (DCE) was designed and administered. Specifically, the way it was done resulted in each respondent "seeing" each disease ONCE. Additionally, there were NO common designs or common choice sets across people. So, one problem is that there will be no variability in choices WITHIN persons for each disease. It simply cannot be the case that individuals exhibit equal consistency in choices with respect to all diseases. Instead, we should expect differences in choice consistency for different diseases, especially in so far as different people will have different experiences with, knowledge of and aversion to different diseases.

RESPONSE: Our choice set design problem involves a very unusual but binding constraint that stems from the need to specify *illness profiles* for each individual's remaining lifespan. This remaining lifespan depends unavoidably on age and gender. If our designs could have been independent of age and gender (as they typically are in most marketing research involving choices among products), we would certainly have employed the state of the art in the design of unconstrained discrete choice experiments. Given this constraint, however, we remain confident that we chose the best strategy available, at least based on any reasonable benefit-cost criterion.

Illnesses as brands: Although we do so in other papers that have since grown out of this project, we do not use the available disease labels ("brands") in this paper (and we likewise do not use the details about severe versus moderate pain and disability, or about hospitalization or surgery). These other attributes were *orthogonal* to the main health state durations and thus we treat them as noise in this analysis. Yes, they may contribute to the error variances, but they are *uncorrelated* with the individual's durations in each of the four health states, their age, or their income. These durations and net income are the only variables incorporated in the estimating specification in this paper. Orthogonality means that the omission of the other attributes from the systematic portion of the model is unlikely to create much bias in the average marginal utility estimates for this model.

The variety of illness labels used in the ten illness profiles in the five different choice sets mirrors the discussion early in the survey that there are many major illnesses from which people suffer and die. We sought adequate coverage for each person of most major illnesses, and a number of specifically for policy-relevant illnesses. We explore elsewhere (Cameron, DeShazo and Johnson, 2008, listed below) how the individual salience of different illness labels affects WTP, noting that it is possible (in our experimentally generated illness profiles) for the identical profile to bear two different illness labels.

Non-constant variances: It is of course entirely possible to define variance components, unique to each choice set (or each risk-reduction program), according to which diseases are included among the two which are mentioned in each set. We have explored models with the error variance normalized to unity for heart disease, and multiplicative terms for each of the eleven other illness labels, switched on or off according to whether an illness profile bearing that label is involved in each choice. Estimation of this model, of course, means shifting to general nonlinear function-optimizing software (we use Matlab).

Impossibility of using just a small number of goods: It is one thing to employ these brand-based scale factor generalizations in a context where there are just five or six products which are identical within each brand. In that case, the parameter space remains of manageable size when "brands" are included as sources of heterogeneity in both the systematic and stochastic portions of the model. Unfortunately, it is not possible to adequately describe the full range of future illness profiles for people of different current ages and genders with just a small number of "brand" indicators. In addition to the twelve distinct randomly assigned illness labels, each illness profile differs in the allocation of remaining life-years across four different health states. There are only four basic variables in our specification: net income, and discounted sick-years, recovered/remission years, and lost life-years. However, the data dictate a flexible translog-type functional form and heterogeneity with respect to current age is also indicated leading to a fourteen-parameter model as the most parsimonious specification.

To keep the parameter space manageable as we conducted preliminary explorations of the need for heteroscedastic errors, we used a simple five-parameter specification that is similar to Model 1 in Table 1 of the paper (except that we use a quadratic form in net income and the three health-state duration variables entered in linear form, rather than logarithmic form). If there is mischief in the error term, it is often most pronounced when the systematic portion of the model is underspecified in some way.

Heteroscedastic models: We have allowed the scale of the error term to differ by illness label (brand) as the referees suggest. The error dispersion parameter is normalized at unity for the a heart disease illness profile. The logarithm of the error dispersion was then allowed to vary systematically with a set of eleven other disease indicator variables, each of which was switched on if that disease was used as a label for the illness profile in question. These models converged readily.

For eleven extra parameters, the maximized value of the log likelihood improves by less than eight points, which suggests that the heteroscedasticity parameters are not jointly significant. Only one individual parameter comes remotely close to statistical significance (i.e. the coefficient for breast cancer has an asymptotic t-test statistic of -1.59). The coefficients of the logit index for the heteroscedastic model average about 1.27 times the magnitude of the coefficients from the homoscedastic model, with overlapping confidence intervals.

	Homo-scedastic	Hetero-scedastic
Linear term in net income	4.53	5.62

Quadratic term in net income	-1.77	-2.17
Sick-years term	-8.81	-9.19
Recovered-years term	-8.23	-12.3
Lost life-years term	-8.38	-11.4

Of course, logit coefficients are known only up to a scale factor that reflects the error dispersion and the unitary scale factor in the homoscedastic model is not the same as the varying scale factor in the heteroscedastic model. The full heteroscedastic model is

Variable:	Coef.	t-test
Linear term in net income	5.6193	2.2187
Quadratic term in net income	-2.1723	-2.0167
Sick-years term	-9.189	-1.8908
Recovered/remission years	-12.3475	-1.7074
Lost life-years term	-11.3681	-2.1692
Dispersion shifters:		
Breast cancer	-0.47303	-1.5916
Prostate cancer	0.33268	0.75322
Colon cancer	-0.048578	-0.14903
Lung cancer	-0.13283	-0.4227
Skin cancer	0.34089	0.8877
Heart attack	0.1457	0.44415
Stroke	0.099084	0.29434
Respiratory disease	0.59716	1.366
Traffic accident	0.57021	1.3912
Diabetes	0.23423	0.66544
Alzheimer's disease	0.009134	0.029036

Richer specifications? Following another suggestion by these referees ("in so far as different people will have different experiences with, knowledge of and aversion to different diseases"), we also attempted a model that included not only indicators for each disease label, but distinct terms for the individual's subjective risk of each type of disease rated on a -2 to +2 scale. Unfortunately, this model with 11+12=23 dispersion shifters could not be coaxed to convergence.

No evidence so far. Had we seen any dramatic results in the form of individually statistically significant dispersion shifters in this simple model, we might have been moved to complicate the full 14-parameter working specification in our general paper by introducing heteroscedastic models. It is still possible, of course, that heteroscedasticity by disease label could emerge as we shift to more flexible specifications for the systematic portion of the model (i.e. a translog-type specification with quadratic age shifters on three coefficients). However, it seems more common that heteroscedasticity problems are *alleviated* by a more general specification for the systematic portion of a model. There is certainly no evidence of dramatic changes in any of our baseline results as we explore a wide variety of additional sources of heterogeneity in preferences in the

list of related papers provided below. These scale-factor results for the simple specification above led us to believe that heteroscedasticity in this case is probably not a first-order sort of problem to be accommodated.

Perhaps another dissertation? To incorporate a comprehensive assessment of potential heteroscedasticity into the present paper would greatly increase its length and make it a very different paper. We already consider the effects of illness labels as they enter the systematic portion of the model in a separate stand-alone paper not yet submitted anywhere for publication, since it has been awaiting the outcome for this paper. (Cameron, DeShazo and Johnson, "Willingness to Pay for Health Risk Reductions: Differences by Type of Illness"—presented at three different conferences since 2008.) In that paper, we can certainly consider further illness-label effects on the error dispersion as well as the systematic part of the model. However, if our considerable past experience with modeling choice consistency is any guide, it will also be difficult to include 22 new parameters in both the systematic and stochastic portions of the much bigger model. The parameter space becomes very large and identification becomes difficult, even though the questions remain interesting. Our choice context cannot be as simple or uncluttered a problem as asking subjects to choose among six snack foods to be eaten now or one or two weeks from now, where the only measured heterogeneity is an 11-point Likert scale for how much the subject likes each type of snack. Our model has to be more complex than this, and its complexity pushes the limits of parameter space given the size of our sample and the necessary heterogeneity in our scenarios.

Indeed, in describing the piloting of the study, the article states the importance respondents placed on considering real diseases. Without any common choice sets or a common design, one effectively confounds differences between people with differences in the choice sets that they received. This makes it very difficult to untangle differences in preferences and/or differences in choice consistency or error variability. In turn, this exposes the authors to the criticism that their results are likely to be biased by failure to take differences in error variability into account. The latter, of course, is due to the fact that fixed or random effects choice models confound parameter estimates (and distributions of such estimates) with the distribution of error variances (or their inverse, "scale").

RESPONSE: We now mention explicitly in the paper that the parameter estimates in our models *may* be biased by our assumption of a common error variance. Yes, there are differences in people and there are differences in choice sets. However these differences are *independent*. The different choice sets are randomly assigned, conditional on the age and gender of the individual, as required by the theory. The mixture of spell lengths for latencies, sick-years, recovered/remission years and lost life-years, and the label assigned to that illness, is completely randomized. Since we have a representative sample of U.S. adults between the ages of 25 and 93, we focus in this paper on the task of estimating preferences for a *representative* individual. Age and income are essential variables because of the need for (1) plausible illness profiles over the entirety of individual's future life-span and (2) the pattern of obligation to bear the costs of risk reductions only when the individual is neither sick nor dead.

Fixed effects: The choice of fixed or random effects models makes a minimal difference in the parameter estimates, which is to be expected since the attribute levels are randomized. By construction, illness profiles vary systematically only with age (and gender, minimally, except for the breast and prostate cancer labels). This same age variable is explicitly controlled for in the model (i.e. it is not "unobserved heterogeneity").

It is now well-known that if one blocks a choice experiment into versions, there will be version effects. This is the ultimate form of that effect, whereby each person receives a different version. There are much better ways to design experiments like this to avoid these types of problems, such as using a balanced incomplete block design or similar design to assign each person a particular set of diseases and/or assigning them to one pair of diseases and asking all questions about that pair. Naturally, this requires a different way of sampling and implementation, but we do this commonly in our research centre.

RESPONSE: Perhaps I misunderstand, but this comment again seems to overlook the fact that eligible illness profiles must be unique to each age/gender combination and we were pushing the limits of survey duration to ask each individual even just five choice questions. In most *marketing* applications, of course, the range of possible alternatives can typically be treated as independent of the individual's age and gender, so the researcher has great latitude in blocking out subsets of the sample to receive different designs. But that is not the case here. One cannot specify the same future illness profile for two people of different ages because a different number of future life-years must be apportioned into latency years, sick-years, recovered/remission years, and lost life-years. Male and female life expectancies also differ at each age, so the gender dimension is likewise unavoidable.

The uniqueness of illness profiles by age and gender is essential and it represents a significant constraint on what this survey—or indeed any other survey about realistic prospective illness profiles—might ever consider for choice set designs.

Version effects: We certainly agree that it is difficult to decide which set of estimates to adopt when there are "version effects" across a number of different survey designs. Again, it is important to realize that each illness profile needs to match the individual's age and gender. Thus we could use no fewer than about 135 different "versions" of the survey in our study. (Table C-2 on page C-16 of Appendix C shows that amongst our 1801 respondents, the number of people in each of these 135 different age/gender bins is only rarely greater than 25 and typically much smaller.) We sought instead to minimize the dependence of our WTP estimates on which "version" of the choice scenarios we used. We seek instead to find the *average* estimates across *the widest possible range* of versions. This seemed to be the only tractable solution, since it was unlikely we would have the precision to identify version effects *within* any individual age/gender group, given the small number of observations in each bin.

2. A second, related issue relates to the discount rate. The way that the experiment is implemented and the analysis done, it may well be that the effect of the discount rate that they report is due to a failure to take error variability differences into account. A recent paper by Salisbury & Feinberg in the journal *Marketing Science* shows that much of the work in intertemporal discounting needs to be reconsidered due to failure to take error variance differences into account, and when they are taken into account, there is no future discounting apparent.

RESPONSE: Actually, the Salisbury and Feinberg (2010) paper seems to be silent on intertemporal choice:

> "Note that this phenomenon is distinct from the primary focus of the literature on intertemporal choice (e.g. choosing between two alternatives whose outcomes are experienced at different times)." (page 3)

This comment is also somewhat perplexing because it seems to imply that this referee thinks we are estimating discount rates from people's choices. Discounting is a maintained hypothesis. Again, the typical inter-temporal marketing choice contexts with their shorter time horizons (i.e. a few weeks at most) may have misled these reviewers. Given the very long time horizons relevant to our choice scenarios (i.e. the rest of the individual's life), it seems *implausible* that discounting does not happen. A discounted expected utility framework seems the obvious and simple place to start for a conceptual model that is as new as the one in this paper. This is not to say that future surveys (or even future papers using our own data) will not focus on this issue specifically, but we do not intend to make it the centerpiece of this paper. Our goal is to explain our basic theoretical framework and to demonstrate its conceptual and empirical utility.

Time delay: Our context also differs from the one explored in Salisbury and Feinberg (2010) because our respondents are asked to make all their stated choices for "consumption" *starting now*. Each independent choice scenario involves enrollment in a health risk reduction program that commences now and will have to be sustained over time to produce a reduction in the odds of suffering a specified illness profile. Consumption of the good in question (the annual diagnostic testing program) begins now, not at some point in the distant future, such as "only when symptoms would finally become apparent." That would be too late. These are prevention programs, not treatment programs. It is true that the characteristics of each illness profile include an expected future time at which symptoms would appear, a duration of sick-time, a duration of recovered/remission time, and a number of lost life-years. These illness profiles are heterogeneous, as they must be if we are to estimate marginal willingness to pay for avoided sick-years, recovered time, or lost life-years. However, it is not these health states themselves which are being "consumed." Rather, the choices involve decisions to consume annual risk-reducing diagnostic tests, starting right now in all cases.

Discounting: The effects of different (common) exogenous discount rate assumptions during our sensitivity analyses seem to have little to do with whether our specifications

are heteroscedastic. We simply report the differences in our estimates and age profiles for WTP as a function of three different assumptions concerning discount rates.

That said, we do have another project in the wings which uses inter-temporal financial choices by a *different* sample of respondents from the same consumer panel as a potential source for information about individual-specific discount rates. That model can, in principle, be employed to estimate fitted individual discount rates in conjunction with individual-specific observable characteristics in the sample used for the model in the present paper. That estimator is rather complex (i.e. highly nonlinear in parameters and requiring a tailor-made likelihood function defined across two different samples). Furthermore, a wide array of additional individual-specific characteristics needs to be brought into play to predict individual discount rates with any specificity. The individual-specific discounting model, if employed, needs to be defined, discussed, and defended. We decided not to pursue that enhancement in the current paper. Instead, we wish to keep the focus on the basic simple model and its distinct advantages over existing approaches to valuation of health risk reductions. The assumption of a common fixed discount rate is of course not uncommon. Anticipating questions about the effects of our discounting assumptions, of course, we carefully addressed their influence via sensitivity analyses.

I am sorry that our lack of clairvoyance back in 2003 made it impossible for us to incorporate the 2010 findings of Salisbury and Feinberg into our survey design or the analysis described in our manuscript and its revisions submitted to the AER in 2005, 2008, or 2009. About the best that can be done as of this point is to cite this more recent work carefully and to encourage anyone who attempts to replicate our study to pay very close attention to these important new contributions to the literature. However, we need to be very clear that even if these new insights had been know at the time our survey was developed, the need for no fewer than 135 different versions of the survey would have severely limited our ability to implement any of their strategies.

It is a common theme throughout the new comments of referees #4 and #5 that we should be paying more attention, in this paper, to choice consistency and the random component in our models. In their 2010 paper entitled “Alleviating the Constant Stochastic Variance…,” Salisbury and Feinberg state:

> In this article, although we hew to the simpler and more standard term “error,” a core concept is that put forward by Louviere et al. (2002), who sought to “dissect” the random or “unobserved” component of utility and who suggested numerous dimensions across which the variance of this unobserved utility could vary.

Our familiarity with the scale issue: We were two of the “et al.” authors of that Louviere paper. It is also rewarding to note that Salisbury and Feinberg, in their conclusions, refer to “As detailed by Adamowicz et al. (2008), different contexts can evoke distinct choice processes and strategies, which may entail differing degrees of unobserved variability.” One of us is among the coauthors on that paper as well. This should confirm that we are certainly aware of concerns

associated with the variance of the error component. We have simply been unable to get much traction along this dimension, empirically, in our work with these data so far.

Once we have published our basic approach to the modeling of preferences with respect to health risk reductions, it is our intention to continue mining these data for other interesting empirical regularities. Although nothing has shown up in the most obvious and simple models that allow for heterogeneity in the dispersion parameter in this study, something could potentially still turn up somewhere if we dig deeper and longer. Indeed, we might mimic and extend some of the models used in the paper entitled "Designing choice sets for stated preference methods: The effects of complexity on choice consistency" published in the *Journal of Environmental Economics and Management*, July 2002, which was one product of a dissertation at UCLA co-chaired by the two of us. (That paper has now been cited 81 times in the ISI Web of Knowledge, although it was unfortunately overlooked by Salisbury and Feinberg, probably because it appeared in the environmental, rather than the marketing literature.)

Scale-related research with these data: Our data from this study have already been the basis for a paper directed to the choice modeling audience, relevant to the issue of the scale factor, and currently under review. "Differential attention to attributes in utility-theoretic choice models" provides a theoretical model based on optimization in the allocation of attention. In this paper, which has been evolving since the 2002 Berkeley Invitational Choice Symposium, we reveal how attention levels, if uniform across attributes and different across people, can manifest themselves as scale differences. The more interesting outcomes, however, stem from cases when attention is unequal across attributes.

We did not send this scale-related paper to the *AER*, even though it represents an empirically viable analog of the 2006 experimental paper in the *AER* by Gabaix, Laibson, Moloche and Weinberg. We assumed automatically that it would be more suitable for a cross-cutting audience of choice modelers and therefore a more specialized field journal, where it has now been revised and resubmitted. We suspected that these issues would be viewed as a second-order concern relative to the main point of this paper for a general audience more interested in health valuation—a simple and streamlined explanation of a wholesale change in the way we conceptualize the valuation of health risk reductions.

3. We were surprised by the fact that they asked lots of questions about diseases and associated risks, but did not use them in their analyses. For example, it seems reasonable to assume that people who are more concerned about the risks posed by certain diseases would respond either to disease or risk more strongly and/or consistently. At a minimum, answers to such questions could have been used to explain differences in individuals.

RESPONSE: We do indeed make use of the many available individual attributes—orthogonal to the choice set attributes—to assess the theoretical construct validity of our

estimated models. During the five years since the submission of the first version of this manuscript, we have actually had plenty of time to explore many extensions and elaborations of the basic model. None of these generalizations produces any big surprises, which is very reassuring. We have merely been adding to the flexibility of the model to accommodate various different types of heterogeneity.

Example: One such paper is entitled "The effect of health status on willingness to pay for morbidity and mortality risk reductions." We control for same-illness and other-illness prior morbidity, and for same-illness and other-illness subjective risks, and even for county-level same-illness historical mortality rates. The effects of this heterogeneity are statistically significant and plausible in their direction and in their effects on WTP. That paper has existed in draft form since 2005 and has been awaiting the outcome of this review process. It would have been straightforward simply to incorporate some of those results into this paper if any of Referees #1, #2, or even #3 had requested this extension.

4. We were disappointed that they dismissed models that are more sophisticated than fixed effects. Holding aside the concerns over confounding each person's unique choice sets with differences in the person, one could have used the recent developments in scale-adjusted latent class models by Magdison and Vermunt to at least try to identify groups that differ simultaneously in error variability and preference variability.

RESPONSE: We have been using these data for years, including in teaching examples to illustrate latent class models and random parameters models. Of course, one is always curious to see what happens when heterogeneity is captured in different ways, and we tried all the usual extensions early on. There are a couple of obvious possibilities (1) assume just a few sets of underlying preferences (as in a latent class model) or (2) treat each basic preference parameter as random (with an arbitrary functional form assumed for its distribution).

Age effects and random parameters models: The dependence of willingness-to-pay specifically upon age is one of the most pressing policy questions. It is *unhelpful,* in that regard, to subsume preference heterogeneity under the assumption that the key parameters are random. As we point out in Section 5 of Appendix E produced for Referee #3, converting our basic parameters to random parameters merely absorbs the heterogeneity in ages and leaves us with very little that is interesting to say, except for the observation that "there is statistically significant dispersion in parameters across the sample." We do not simply "dismiss" models that are more sophisticated than fixed effects (where even fixed effects logit models make precious little difference to the parameters of the model, given the random assignment of attributes). We have actually pondered random parameters models very carefully over the years and found them to be wanting, relative to the models we use in the paper.

Latent class models: Likewise, we do not merely dismiss latent class models either. We have explored those specifications and found that none of the obvious latent class models (at least those with few enough classes to converge) do nearly as well at explaining preference heterogeneity as we can do with our systematically varying parameter

models—when the point is to explore the sources of such heterogeneity. This is a very rich data set. Luckily, due to the orthogonality in the design, it is possible to entertain different types of heterogeneity, one at a time, in papers of manageable length. We regularly take advantage of the fact that the attributes of our illness profiles are all independent of individual characteristics (except for age, which is explicitly captured in all of our models with six different interaction terms).

If there is a specific Magidson and Vermunt paper that we should somehow incorporate among our references, and acknowledge at some point in the paper, we would be happy to do so. Readers may wish to know how latent class models have been appropriate in other contexts. The candidates appear to be:

Latent class analysis with sampling weights - A maximum-likelihood approach
SOCIOLOGICAL METHODS & RESEARCH Volume: 36 Issue: 1 Pages: 87-111 Published: AUG 2007

Current issues and a 'wish list' for conjoint analysis
APPLIED STOCHASTIC MODELS IN BUSINESS AND INDUSTRY Volume: 21 Issue: 4-5 Pages: 327-328 Published: JUL-OCT 2005

Latent class models for classification
COMPUTATIONAL STATISTICS & DATA ANALYSIS Volume: 41 Issue: 3-4 Pages: 531-537 Published: JAN 28 2003

Latent class factor and cluster models, bi-plots, and related graphical displays
SOCIOLOGICAL METHODOLOGY 2001, VOL 31 Book Series: SOCIOLOGICAL METHODOLOGY Volume: 31 Pages: 223-264 Published: 2001

The last two seem to have been cited 23 and 47 times, respectively, whereas there have been no citations to the first two yet. I'm not sure which one of these papers the referees had in mind.

Our conclusion is that there is much to recommend on the theoretical development side, but far less than meets the eye on the stated preference and modelling side. It also is unclear whether the issues noted above can be sufficiently dealt with and/or reconciled given that the data are what they are. Interacting variables that describe the respondents' views about particular diseases with the choice data may serve to model disease-related differences in choice consistency but given the lack of within-respondent variation this is a second-best approach which assumes that such effects are constant across respondents.

RESPONSE: We are grateful that these referees recognize that our theoretical approach to modeling the valuation of health risk reductions is new, appropriate, and apparently sound. But we are also confident that a good part of the reason that they feel there is "less than meets the eye on the stated preference and modeling side" stems from our failure to make it clear in the paper that there can be no fewer than about 135 different versions of

the survey, given that the same illness profile can never be used with more than one gender/age combination.

Stated preferences: While many mainstream economists remain skeptical of stated preference research, there is truly no kind of revealed preference data that would permit the insights we derive in this study. However, we perceive that these referees are not opposed to stated preference research per se. They simply would have liked us to be able to use a blocked design in our study, typical of high-quality marketing studies, which would have allowed us to develop simpler and more clear-cut models for scale differences in the error terms.

Unfortunately, these simple blocked designs are not really possible, given the small numbers of respondents in each gender/age bin. Salisbury and Feinberg (2010) appear to have used four groups of 24-28 participants, vastly exceeding our maximum number of respondents in any single gender/age bin in our study. Furthermore, their choice scenarios and models can be much simpler, since individuals merely choose between six standard snack foods (apparently without even being required to pay for their choices) and the only respondent-specific data is the individual's rated like/dislike for each type of snack food in the set of six.

Complementarity: Given the simple and clear framework for their study, it is not really surprising that Salisbury and Feinberg are able to identify the interesting systematic scale effects that they find. Their study makes a significant methodological contribution, but one could lodge the criticism that the choice context is contrived and oversimplified, the preference function is completely ad hoc, and the application in question is not policy-relevant at all. However, these valid criticisms do not detract from the significant innovations they offer, just as we hope that the constraints on our experimental design options, and the potential for further relaxation of the stochastic assumptions than we have already explored, will not detract from the significant innovations we make in this paper. These include the richness of our choice scenarios, our utility-theoretic preference function, and the policy-relevance of the issue we address. Realistic choices among costly health-risk reduction programs targeted at a broad spectrum of morbidity and mortality risks are unavoidably harder to capture than choices among six snack foods in a laboratory setting. These two veins of research should be considered to be complementary. Each has a comparative advantage in what it offers, but neither will satisfy all possible consumers of the research.

Despite the possibility of bias due to things that might be going on with the scale factor, it is important not to lose sight of the fact that our resulting generalized WTP estimates for mortality and morbidity risk reductions are completely consistent, in their appropriate special case, with numerous standard WTP estimates for mortality risk reductions based on revealed preference methods such as wage-risk models. At this baseline, then, there is no evidence to suggest we are finding anything that is biased to the point of being fundamentally implausible or severely misleading. There is thus no *a priori* reason to doubt the potential value of what our model can produce over the broader unexplored domain where no revealed preference data exist.

Other work: In our third-round submission to the AER, perhaps we should have included all of the following additional papers, *based on the same data set*, as additional appendices. On several of the more peripheral papers, we recently gave up waiting for the AER and went ahead and submitted them anyway.

"The effect of health status on willingness to pay for morbidity and mortality risk reductions," J.R. DeShazo and Trudy Ann Cameron (pending submission)
"Two types of age effects in the demand for reductions in mortality risks with differing latencies," J.R. DeShazo and Trudy Ann Cameron (pending submission)
"Willingness to pay for health risk reductions: differences by type of illness," Trudy Ann Cameron, J.R. DeShazo, and Erica Johnson (pending submission)
"Differential attention to attributes in a utility-theoretic discrete choice model," Trudy Ann Cameron and J.R. DeShazo (under review UPDATE: published in *Journal of Choice Modelling*.)
"Demand for health risk reductions: a cross-national comparison between the U.S. and Canada, Trudy Ann Cameron, J.R. DeShazo, and Peter Stiffler (under review UPDATE: published in *Journal of Risk and Uncertainty*)
"'Scenario adjustment' in stated preference research," Trudy Ann Cameron, J.R. DeShazo, and Erica Johnson (under review UPDATE: *published in Journal of Choice Modelling*)
"Subjective choice difficulty in stated preference surveys," Eric Duquette, Trudy Ann Cameron, and J.R. DeShazo (under review)
"The effect of children on adult demands for health-risk reductions," Trudy Ann Cameron, J.R. DeShazo, and Erica Johnson (forthcoming, *Journal of Health Economics*)

Every one of these papers in the above list has now been presented at conferences and workshops, and multiple times in some cases, given the series of Ph.D. students we have involved in the project. Across all of these outings, we have received a lot of feedback from discussants and session participants which has helped us address a wide variety of potential issues. The process is becoming asymptotic.

In response to the referees' final point about the disease information: we note that the paper that focuses on heterogeneity in WTP for risk reductions by type of disease (controlling for subjective risks, perceived controllability, and a host of other factors) formed the third paper of a 2008 dissertation by Erica Johnson. It has also been awaiting a decision on this paper before it could be sent anywhere. We are grateful that these referees consider Erica's other extensions to our basic model to be a worthy endeavor.

7.2.4 Editor's final decision at the *AER*

"I can now give you a decision on your *AER* resubmission with DeShazo, "A Generalized Empirical Model of Demand for Health Risk Reductions." I apologize for the delay, but this has been a difficult manuscript to deal with.

I do not have a reply on the resubmission from Referee I. I did not send it back to Referee II.

Referee III had the most fundamental objections to your paper. That referee read your resubmission materials and replies, and decided to withdraw from further review of your paper. The referee said that he/she was not convinced of your replies to his/her original report and continued to have major concerns about the design of your experiments. However, the referee also said that he/she did not want to be the one to recommend against publication, given the wide circulation and attention your paper has received."

> *[Note: This statement conflicts somewhat with what Referee III reported to us himself, having revealed his identity: "The deadline for submitting my review on the revised AER manuscript came and went last week. I started wondering why I had procrastinated doing the review, even though it has appeared in my Outlook reminder list every morning for a couple of months. I then called XXX and withdrew as a reviewer. I told him I felt too conflicted and exposed to give him a completely objective recommendation. He asked me to suggest a couple of alternative reviewers, but I sincerely hope he doesn't bring yet another player into this endless process." It is not clear that the AER editor remembered that this reviewer's identity was known to us. It is not possible to know whose version of their transaction was correct.]*

"Recognizing that the design issues of your experiments are the major issue, I searched for another referee who is expert at these methods. I located a senior, respected scholar who was strongly recommended to me by more than one person. This individual brought a colleague of his/hers into the review and wrote the attached report. The report thinks that health economics needs to reform and replace existing methods and that you have provided a solid theoretical approach but thinks that there are issues with implementation. The report lists four concerns, the first and longest of which relates closely to one of the key issues which concerned Referee III. The report also concludes with the concern that, given your data, the issues the report raises cannot be fully dealt with."

> *[Note: Referee III mistakenly thought we had <u>too few</u> versions of our survey instrument (i.e. just one), while referees IV and V thought we had <u>too many</u>, since they didn't realize that it is not possible to give the identical survey instrument to people of different ages or different genders. Thus their concerns on this matter are thus opposite, and both are misguided.]*

"I have looked again at your paper but, necessarily, must rely heavily on the referee reports in this case. I am afraid that I cannot proceed any further with your paper, given these reports. I have one favorable referee (Referee II) and three unfavorable ones. Referees III and IV share concerns. As I mentioned to you in my last letter, we try hard at the *AER* to stop the process of revision and resubmission unless all major issues appear to be resolved or resolvable, and that is not the case here. Indeed, I was quite reluctant to ask for a revision even in my last letter to you. While I am sure that you have responses to Referee IV, and perhaps more to Referee III, I am afraid that I must stop the process now.

The referees have made clear that your study has much value to it in many respects and that it has already been widely circulated and has made an impact. Your theoretical argument in your paper seems very strong to me, too. Given this, I think that a lower-ranked journal which is willing to suspend some doubt about your methods may be likely to accept and publish your

paper. One journal that you may wish to consider for this is *AEJ: Applied Economics.* The advantage of *AEJ: Applied* is that you have the option of asking them to supply all the AER reports and materials to them. This greatly speeds up the process of their decision-making and you would likely receive a quick answer from them. If you choose to submit to them, please indicate to them in your Cover Letter that they should contact me to obtain all AER materials, and also check the box on their submission page indicating that you would like to have them receive the AER reports."

> *[NOTE: I asked the editor whether "suspend some doubt about your methods" was a reference to the stated preference nature of our data, but I received no reply.]*

7.3 Submissions to the *American Economic Journal: Policy*

We also thank our editor at the *American Economic Journal: Economic Policy.* Although the *AER* editor recommended we submit the paper to the *AEJ: Applied Economics* after his rejection, the *AEJ: Economic Policy* is the journal that explicitly covers "public economics; urban and regional economics; public policy aspects of health, education, welfare and political institutions; law and economics; economic regulation; and environmental and natural resource economics." Given that our paper is intended to improve public policy decisions with respect to health, safety, and environmental regulations, this seemed to be the most appropriate alternative.

However, the *AEJ:Policy* editor desk-rejected the paper based on the file transmitted from the *AER*, but made no mention of looking at our rebuttals to the fourth and fifth reviewers, which were submitted along with our question about *whether* the paper could be evaluated as is, or *whether* it would need to go out for review. Her response to our submission is reproduced in the following section.

7.3.1 Response to transfer of *AER* file to the *AEJ:Economic Policy*

"This paper was a transfer from the *AER*. I have carefully studied the paper and the review materials from the *AER* (the referee reports on the three drafts of the manuscript, the three decision letters by [the *AER* editor], and your responses to the first two rounds of referee reports).

I appreciate that you are asking for, and expecting, an up or down decision on the manuscript. Unfortunately, after reading through the reports and letters from the *AER*, I am not convinced that the paper is appropriate for the *AEJ Policy*.

My decision is partly based on my view that the paper is a better fit for a more specialized journal and readership. While the use of these valuation methods is an important input to a wide range of cost / benefit type analyses, the analysis in this paper seems better suited to those scholars who are working more closely on these problems. Further, I found several of the issues around the survey design and methodology raised by *AER* Referees XX and YY [sic] to be unresolved in the exchange back and forth. Since my own expertise is far from this area, I simply am not in a position to move forward with the paper without further consulting with my editorial board.

Consequently, I am rejecting the paper.

I would recommend trying the *Journal of Risk and Uncertainty*."

7.3.2 Response to paper on illness-specific valuation

Given that one of the complaints by the final *AER* referees concerned the fact that we had made no use of the (orthogonal) illness names in our analysis, and our random guess that this could be one of the unspecified "several" unresolved issues mentioned above, we decided to submit our related paper that explores the influence of illness names to the *AEJ:Policy*. The same editor's reply was as follows:

"I have read your paper and find it quite interesting and relevant for public policy analysis. The ability to identify and estimate illness-specific valuations is novel, and brings to mind several follow up issues such as what this reveals about preferences, perceptions of risk, and so on. Given the prevalence of cost benefit analysis in many aspects of public policy, this type of analysis is of particular interest.

Unfortunately, I have concluded that it is not appropriate for the *AEJ Policy*. While your topic is likely to be of great interest to specialists, it is not of sufficient breadth for our journal. I would recommend the paper would be a better fit for a more specialized journal such as *Journal of Health Economics*, *Journal of Risk and Uncertainty*, or possibly, the *Journal of Public Economics*."

> *[NOTE: The August 2011 issue of this same journal includes papers on topics as narrow as the length of postpartum hospital stays, exposure to environmental tobacco smoke, and the effect of school start time on the academic achievement of adolescents, so we remain mystified about how our model and comprehensive estimates for the general population of the U.S. of demand for policies to reduce risks from twelve of the most common life-threatening illnesses is "not of sufficient breadth." Perhaps the paper is just too complex and serious and not "catchy" enough to appeal to an audience of casual readers.]*

7.4 Reflections on our submissions to general-interest journals

As unusual as it is to reveal the machinations of the review process with respect to a research paper, there is a pressing need for full disclosure when the results of a research initiative are intended to be useful to the policy-making process. We wish to reassure the potential consumers of our findings that none of the issues raised in the peer review process for this research have been suppressed. The comments of our various editors and reviewers raised concerns that certainly might have been shared by other readers of earlier versions of our paper, especially prior to the numerous revisions incorporated into the current manuscript and before this supplementary document containing information and auxiliary analyses became available.

Given the substantial public funds that have been dedicated to this research, we had certainly hoped to publish the research in a general-interest outlet, to maximize our audience of non-specialist economists. We thus made this concerted effort, over the course of more than five years, to place this research in a mainstream general-interest economics journal. We felt this effort was necessary, since the willingness-to-pay measures our research provides are widely relevant to all types of government policies and regulations concerning health, safety, or environmental quality.

What is our own critical assessment of our review and revision process? We encountered two types of reviewers at these journals. One group of non-specialists focused on the potential

for the paper to appeal to a general audience with little prior investment in the issues being addressed or the types of data and methods being used. The other group, consisting of highly experienced specialists, pushed us to address many important technical issues at the frontier of stated preference and conjoint choice research. Our efforts to produce a single paper that was simultaneously "conversationally general" and "technically rigorous" were ultimately unsuccessful.

Our initial submission to the *JPE* was undoubtedly premature, since we had not stepped back to consider just how little of the background for the paper would be common knowledge among other economists. Then the review process at the *AER* hinged at one point on the opinion of tie-breaking (and incredibly thorough) Referee III, who decided to bail out of the review process without providing us with any feedback on our responses to his requests. Referees IV and V wanted/needed even more discussion of other alternative methods and models, but we were not permitted to engage them on their points. Our submission to the *AEJ:Policy* was rejected because the *AER* rejected it. (Both editors concede their lack of expertise in the relevant research areas.) Our follow-up submission to the *AEJ:Policy* of one of the applications of our model and data—our "illnesses" paper—was acknowledged as being interesting and policy-relevant, but was desk-rejected as being too specialized.

In 2010, we gave up on our effort to craft a paper for a general-interest economics journal that was simultaneously very accessible for non-specialists and completely satisfying to specialists, all within forty pages or less. Since a large share of the funding for the project was contributed by the U.S. Environmental Protection Agency, we decided to send the paper to the leading field journal for this subject matter, the *Journal of Environmental Economics and Management*.

7.5 Submission to the *Journal of Environmental Economics and Management*

7.5.1 Editor's Substantive Comments and our responses

Reviewer #1 was quite positive, although he/she notes (i) issues with respect to income effects associated with a major illness and (ii) the distracting effect of referring to a large set of appendices (note – reviewer #1 submitted his/her review before the situation with your technical appendices was resolved and did not see them). I suspect that you will not be able to address the income effects question directly (at least not without going back and re-surveying respondents), but it seems like a point that deserves some explanation.

> RESPONSE: We have been able to explore different assumptions about the effects of lost future income on respondents' WTP. Please see the detailed discussion and results in our response to Reviewer 1. We also discuss the future income assumption on page 16 of the manuscript (including footnote 32), and on page 28-29. Near the top of page 32, we discuss the new fourth column of WTP estimates added to Table 3, based on a model where respondents are assumed to expect zero income while sick, but normal income upon recovery/remission. We also expand upon this comment with the evidence presented in Section 5.6.10 in the online Handbook.

Dealing with point (ii), I would also ask that you find some way to incorporate the information in the technical appendices more seamlessly into the paper.

RESPONSE: We have sought to do this in two ways. First we have reduced the number of times we reference this material by grouping our discussion of issues in a more consolidated fashion. Second, we have combined all of the previous appendices into a single online Handbook which now exceeds 300 pages in length (not including an illustrative example of the individually tailored survey instrument). Researchers who use census data or other government sponsored survey data can of course refer to data documentation provided by the agency that collects the data. Given that our data are original, however, we must provide the analogous data documentation ourselves.

The challenge that we face in writing a journal article is that frequently, when a reader or reviewer wonders if we addressed an issue, our most effective response is to direct them to the online Handbook. At this point, fully fourteen total reviewers (editors and referees at the JPE, AER, and AEJ:Policy journals) have speculated about possible biases in stated preference research and the possible sensitivity of our results to alternative specifications. Over the last six years, we have implemented every suggested test or alternative specification that has been feasible with our data. Some of these suggestions have improved our specifications, but others have had negligible effects. Along the way we have also recognized, on our own, a number of important improvements in our models that have not been suggested by any editor or reviewer. All of this experience needs to be documented somewhere if our results are to be used in policy-making applications.

Unfortunately, several reviewer suggestions over the years have also been well-intentioned but misguided, perhaps because we failed to adequately explain our survey data, our model, or our innovations within the page limits common for journal articles. It is unfortunate that our paper has been unsuccessful at these different general-interest journals in large part due to these misconceptions, many of which we have not been permitted to rebut. We now document all of these adventures in the Handbook.

I'm not expecting you to include all of it (see below), but rather to work it in such that this paper can stand alone without the reader being forced to go to a long set of appendices. Reviewer #1 has a short list of additional minor comments that should be easy to address.

Reviewer #2 (the one who read all six technical appendices…) was also positive, but had a few more comments for you to deal with. In particular, he/she felt that the paper was "incomplete" in a number of dimensions (as though it was a piece that had been cut from a larger project). In some sense, this echoes the sentiments of reviewer #1 (I had a similar reaction upon my initial reading as well). In particular, reviewer #2 would like to see a more frank discussion of what is or will be treated elsewhere, and he/she would like to see material from Appendix E included in the main body of the paper (with much of the rest of the material included in a streamlined fashion along the lines described by reviewer #1).

RESPONSE: We suspect that the feeling that it was "cut from a larger project" comes from our frequent referencing of the previous technical appendices. As we note above, we have reduced the number of references to the different appendices by consolidating these discussions, and the details concerning all other digressions, into an extensive online Handbook designed to support all papers in this series.

We also now explicitly discuss in the Introduction and Results section those topics that we treat elsewhere in other papers. See page 9 and footnote 16 where we emphasize that the *random assignment* of illness names and attributes, subject only to some exclusions for implausibility, allows us to avoid any omitted variables bias when we exclude other characteristics of each alternative in this study. We also mention other papers which generalize our basic model on page 27, in footnote 51.

Section 6 of the online Handbook also provides a detailed inventory of each of the eight other papers we have prepared, based on this data set, including titles, coauthors, abstracts, and additional details about the nature of the sample used and the models employed. This present paper is the "flagship" paper. A simpler and much earlier version of this paper (based on a subset of the data and using far fewer corrections for scenario adjustment/rejection) was the foundation upon which each of the other papers were built. Those manuscripts which are yet unpublished will need to be revisited in light of the more-general model developed in this paper.

Reviewer #2 goes on to make a number of important points (1) – (5). Again, I would not expect that you will be able to make substantive changes that require resurveying respondents (aside from describing potential shortcomings of your analysis), but I do think you will be able to address many of these points directly. Reviewer #2 finishes with a couple of general comments that can be dealt with in the interpretation of your results.

RESPONSE: We have addressed all of Reviewer 2's technical questions. Fortunately, many of the issues raised in comments (1) through (6) were things we had already explored and/or dispensed with, but did not have room to include in the paper. Several of these considerations are now dealt with in detail in the online Handbook, and merely mentioned in passing (or in new footnotes) in the paper.

We have conducted several new data analyses and updated our models, partly in our response to this reviewer and partly as a result of our own further deliberations and explorations concerning the most appropriate specification.

Reviewer 2 also raised issue #7 (about our use of co-payments as an acceptable payment vehicle). We have explained why we chose this approach and that respondents appeared to accept it during our numerous one-on-one pilot sessions with Knowledge Networks panelists at the company's Menlo Park facility . There seemed little else we could do on this point. We are also frankly hard-pressed to think of an alternative payment vehicle for private health-risk reduction choices that would have been any more plausible than the one we used, and the referee does not suggest one.

7.5.2 Reviewer #1 comments and our responses

INTRODUCTORY RESPONSE: We appreciate your thoughtful review. We have sought to address both your major concerns and those contained in your "shorter comments." In particular, we have reduced the number of references to the earlier set of appendices. These appendices have now been consolidated into an extensive online Handbook to accompany our broader study. (We have asked the editor to be sure that this Handbook is available to you. It has been extensively indexed to facilitate look-up of how we have handled tangential issues that may still concern you.) Our reduction in the number of references to technical issues reflects our best effort to prioritize of the issues we believe will be of concern to the average reader. We remain open to any further suggestions you may have about how we might better prioritize these issues and the extent to which we should describe the related work in the paper (as opposed to the extensive annotated inventory we provide in Section 6 of the online Handbook).

The structural random utility choice model is rich with implications and well suited to guide estimation of WTP for changes in probabilities of various illness profiles over the life cycle. One aspect that is potentially troubling, however, is the assumption that "the individual expects to retain approximately their current income in real terms through a major illness" (page 19). For major illnesses such as stroke, lung cancer and Alzheimer's disease (page 10) one might expect major income effects. The assumption might apply to government workers on salary, but the reduction in income would probably be sizable for individuals who work in sales or construction. Including the potential income loss would probably increase the WTP for better life profiles without sickness and reduce the expected difference between the WTP for them and WTP for a profile with a lower probability of sudden death. It would be more consistent with results reported in Tables 2 and 3 where differences are smaller than some might expect.

RESPONSE: We agree that the assumptions we make about individuals' future income levels under different health states merit some sensitivity analysis. On page 32, at the end of the section on WTP as a function of income, we discuss a new column added to Table 3. To explore the effects of different assumed levels of future income, we vary the γ_1 term in our more-general model (which is laid out step-by-step in Section 5.1.1 of the online Handbook). The results of assuming that the individual expects to earn zero income while sick are shown in this new column. For greater detail, Section 5.6.10 of the online Handbook provides parameter estimates and WTP simulations when this "fraction of income earned while sick" is set either to 0.5 or to 0. Tables 5-15 and 5-16 in the Handbook show how these two types of income assumptions affect the estimated utility parameters (since the estimating variables change slightly with these different assumptions) and how these different utility parameters affect WTP for each of the full set of illness profiles mentioned in the paper.

The bottom line is that the effects of zero earnings while sick are less noticeable than one might initially expect. Recall that assumed income expectations during lost life-years remain at zero in either case, so it is only the amount of sick-time that makes the difference. All time-periods enter the model in presented discounted terms, so lost future income due to sickness many years into the future will be extensively discounted. There

is no sick-time in the benchmark sudden-death scenario, so earnings while sick are not relevant in that case, except to the extent that this assumption means we construct somewhat different estimating variables and the utility parameters are therefore slightly different.

We assume that individuals expect their regular income to be restored upon recovery or remission of the illness, over the subsequent time period when the individual remains alive (at least in terms of people's ex ante expectations). We are also careful to point out in the paper (e.g. in the footnotes to Table 3) that WTP to avoid a scenario with just one near-term sick-year simultaneously reflects WTP to avoid of the remainder of the individual's life-span in a recovered or remission state following that illness. This post-illness health state is not considered by respondents to be the same as current health (according to the estimated coefficients on the "recovered-years" terms in the model, which would otherwise be zero). Crucially, WTP to avoid the illness profile with one year of major illness is not simply "WTP to avoid one year of major illness." Instead, it is "WTP to avoid one year of major illness followed by the remainder of life in the post-illness recovered state."

Somehow the author(s) should come to terms with what is in this paper and what is to be found elsewhere. One short appendix is effectively included with this paper to provide an example of a choice set. Five other appendixes (A through E) are referred to about 20 times, however, and are not provided. They probably elaborate on important aspects of the paper. Including them would probably make the paper longer than the typical journal article. Nonetheless, it is distracting to be referred to work elsewhere frequently. If they are going to be used, it seems like the first appendix referred to should be A, not C as on page 3. It is easy to empathize with the challenge of distilling the results of an ambitious project into a "short" article, but decreasing the distraction of multiple references to one appendix or another will be worth the effort. One possibility is to combine the five appendixes into a long working paper with chapters and make it available upon request or online.

RESPONSE: The original five appendices were provided, but somehow were not transmitted to this referee, which certainly would have left him/her with questions (as a non-casual reader).

Yes, we agree. We have reduced the number of times we refer to the supplementary material by about one-third. We have also incorporated these earlier appendices into the online Handbook which we now make available to support the entire suite of papers being produced using the data from this survey. Considerably more material has been generated in the course of responding to this set of referee comments, and also as a result of our own further deliberation about appropriate specifications and fewer exclusion criteria. As a result, this online Handbook now exceeds 300 pages of discussion and alternative or tangential results. To make the online Handbook easier to use, we include a detailed Table of Contents, lists of Tables and Figures, and a very comprehensive index.

Short Comments

Pg. 4, fn. 3: Kochi, Hubbell, and Kramer 2006?

We are not sure to what your question mark refers. Could you please clarify?

Pg. 5 & 37-39: An advantage of the approach is the ability to identify inter-temporal substitutability and complementarity among future health states. Shouldn't the related results be discussed in the conclusions?

Yes, we now discuss this in our conclusions section on pages 37 and 38 (e.g. "Our model allows for substitution across different types of health risks with different time profiles…"; "…the severe prior morbidity that may be associated with many mortality risks…", '…heterogeneous marginal values, which depend upon the mix of health states in an illness profile and the individual's age…').

Pg.6: "Micromort" is introduced here, but never used again.

We prefer to use the term "microrisk" to describe both mortality and morbidity risks, generically. We now clarify this preference on page 4. The term "micromort" no longer appears in the paper.

Fns. 45 & 49: Presumably the separate and related papers by the authors will be cited after the review process.

Yes. Section 6 of the online Handbook now describes in detail the precise territory covered by each of our eight other papers which also rely upon the data from this survey. Some of these have already been published. The published papers rely upon a smaller subset of the data (due to additional exclusion restrictions and fewer scenario adjustment/rejection control variables than we use in this paper). These papers variously combine our U.S. and Canadian samples, explore the influence of household structure on WTP, or explore the determinants of respondent attention to the different attributes of the choice set. The remaining unpublished manuscripts will not be finalized until publication of this current paper, since they will need to be modified to reflect updates to the basic estimating specification (before we can generalize the "basic" model in this paper to consider heterogeneity as a function of illness names, subjective risks and current morbidity, or two different conceptualizations of age).

Pg. 29, line 5: Isn't the "sudden death" profile chosen purposely for comparison to VSL studies as a benchmark rather than chosen "arbitrarily"?

This is true. We rephrase the sentence to remove the world "arbitrary." In our conclusions, we distinguish between the choice of the VSL-like scenario and the arbitrary selection of other illness profiles as illustrations of our model's capabilities.

Pgs. 32-33 & 36-37: The discussion of WTP as a function of income might benefit from updating to include Kniesner et al. *JRiskU* (2010) and Evans and Smith *JRiskU* (2010). The discussion of WTP as a function of age might benefit from updating to include Hammitt and Haninger *JRiskU* (2010) and Blomquist et al. *ResEnergyEcon* (forthcoming).

Thank you for these updated citations. Kniesner et al. (2010) and Evans and Smith (2010) are now cited on page 31, Hammitt and Haninger on pages 33 and 34, and Blomquist et al. on page 34.

Table 4: It is interesting that for a 35 year old the WTP for a microrisk reduction to avoid sudden death at age 65 is about equal to the WTP for a microrisk reduction to avoid sudden death now for a 65 year old. It might be worth discussing this in the context of life cycle models of WTP for microrisk reductions.

Our use of a range of scenario adjustment/rejection controls, and the fact that these controls permit us to retrieve sets of observations that were previously excluded from the analysis a priori, have led to substantial changes in the entries in Table 4. This comment is now obsolete.

Table 4, note a: Can't a negative marginal utility of income be ruled out?

Theoretically, it is reasonable to constrain the marginal utility of net income to be positive, and this could be accomplished using a general function-optimizing algorithm. However, we note that "illness" should convey negative utility overall, relative to the status quo, but it is probably not appropriate to restrict *each* adverse health state to convey negative utility. This is because there can be "fates worse than death." We need to allow for the marginal utility of an additional lost life-year to be potentially positive if the preceding morbidity is bad enough and long enough (as with some cases of Alzheimer's disease, potentially). In our current specification, any simulated negative WTP values are artifacts of the parameters being unconstrained, so we can simply interpret negative WTP amounts as zero. The worst a respondent could do would be to not pick either program. There was no opportunity actually to express a negative WTP. The intuition for treating negative fitted values as zero is much like that used in a Tobit model. See footnote 60 in the paper, and the footnotes to Table 4.

7.5.3 Referee #2's comments and our responses

> INTRODUCTORY RESPONSE: You clearly devoted a great deal of time and effort to your review of our paper. We really appreciate your investment and the revised paper is better for it. We have sought to address all of your major concerns (1-7) as well as those contained in your "minor comments." In two instances we provide new analyses to address your concerns.

Several questions can still be raised, related to this work. A first is where this work has actually taken us. The authors claim that this study represents a model for how to conduct VSL assessments for public policy decision purposes, and here I fear that the answer is not necessarily yes. There are also a number of methodological (and presentational) issues that I come back to below.

I will now first list some methodological issues. The paper could here in my view have done a better job (some issues may be presentational; the high degree of complexity of the study makes it difficult to fully decipher it, at least for me). A main point of mine is to point out how certain issues can be better and more completely addressed, when building on the survey material.

1) It appears from the description that each respondent has been subject to 5 different choice situations, each involving 2 specified choice sets (in addition to the "status quo"). These choices involve rather intricate descriptions of complex hypothetical situations (to exemplify, a disease will be acquired in 10 years, will persist for 5 years and ultimately reduce your expected lifetime by 6 months at that stage; this outcome combination must be compared to another, similarly complicated, set of circumstances; and in both cases there are payments to be made, over the entire future lifetimes which generally differ under different alternatives). It is claimed that the choice sets are "tailored" to the individual in terms of age, preferences, etc. My worry is that these specified sets still represent a rather complex hypothetical choice situation for the respondent. My own experience, from similar stated choice surveys, is that the ability of respondents to reason consistently over such alternatives is rather limited.

> RESPONSE: We agree that it is extremely important to understand, and work within, respondents' cognitive constraints in a conjoint choice experiment. Our efforts to understand respondents' cognitive constraints began with earlier research on how alternative choice set designs affect choice consistency (DeShazo and Fermo, 2002) and a model of how respondents allocate attention across choice sets (Cameron and DeShazo, 2010, where a working paper based on the theory was first prepared in 2002, and a later version with an empirical application has now been published in the *Journal of Choice Modeling*).
>
> We paid particular attention to these issues as we began focus group work for this survey and during the design of the choice sets and their attributes and alternatives. Specifically, concerns about "information overload" guided our selection, definition and presentation of attributes. We restricted each choice set to only two alternatives (relative to the status quo). By framing the attributes as part of an illness "story" (which we call an "illness profile") we sought provide a vernacular and familiar framework for the risk outcomes of interest.

First, we developed each respondent's familiarity with the choice process and the attributes of the objects of choice through our attribute-by-attribute tutorial. Second, in light of individuals' limited ability to process complex risk information, we undertook some simplifications in how we represented illness profiles. Our first simplification is that we ask individuals to make choices as if they faced *only* the one given illness profile for that illness.

Our second simplification is that we do not represent illness profiles as compound probabilistic events. When health researchers consider a concatenation of health states, they often first ask a question like: "What is the probability of experiencing prostate cancer?" Then, conditional upon the type of occurrence at a particular age, health researchers describe the conditional probability of survival. We simplify the representation of a series of conditional events by describing a single probability for that series of health states.

Third, we continue to evaluate choice difficulty even during the survey itself by including a difficulty rating opportunity after each choice scenario in the survey. We found some systematic variability respondents' perceptions of choice difficulty, but these variations do not systematically bias our results for any particular illness profile because of the random order in which named illnesses and illness profiles were paired and displayed both within and across respondents. (Our "choice difficulty" results are discussed below).

Perhaps most importantly, if respondents had been completely overwhelmed by choice set complexity or if they had been entirely inattentive to attribute information, their choices should have become more random and the estimated coefficients on the attributes of the illness profiles might not have been statistically significantly different from zero. Perhaps the greatest indication that we effectively address this concern is that individuals recognized and clearly valued changes in all of the core attributes levels analyzed in this paper; their basic estimated (dis)utility coefficients are all statistically significant and of the anticipated sign, and their implied willingness-to-pay estimates are well in line with those produced by other methods in the special case of "sudden death in the current period." For more details please see Sections 2.7-2.9 in the online Handbook (i.e. Basic tests for theoretical validity, Respondent learning and fatigue, Heuristics and metric recoding).

2) Another issue is possible sequencing bias related to the 5 choice situations facing each individual, which is (as far as I can see) is not addressed.

REPSONSE: We minimized the potential for *any* type of "sequencing bias" by randomizing the order of appearance, across all ten illness profiles (grouped into five pairs) in each survey instrument. Each named health threat has a randomly assigned illness profile, with only a few exclusions for implausibility (e.g. no sudden death from diabetes or Alzheimer's disease). The number of permutations of "eleven illness names taken ten at a time" defines the variety in our ordering of illness names across respondents. The universe of possible randomized illness profiles associated with any given illness also varies with the person's age and gender. If we had used only a few sequences of illness types and illness profiles in our study (or just one survey instrument, as a prior referee assumed), this would indeed be a concern, but in our study, every

survey instrument is virtually unique. This is a particular strength of our survey. Our parameter estimates thus reflect the average across all of the major illnesses and all of the illness profiles represented in our study. Our randomization strategy is now described in more detail on page 9.

However, in response to your concerns, we have now explicitly tested for systematic shifts in each of the estimated utility parameters as a function of the position of the choice set among the five choices faced by each respondent. To keep the number of estimated parameters from getting out of hand, we interact each variable in the estimating specification with a linear term and also a quadratic term in the "deviation of the position number of each choice relative to a designated baseline choice number between 1 and 5." This makes it easy for us to simulate WTP for each illness profile normalized on any given "choice number" by setting the deviations (and hence these interaction terms) to zero. These systematic variations in parameter estimates are presented in Table 5-18 of the online Handbook. (This model also allows each parameter estimate to shift according to the deviation from the median time spent considering each choice set, thereby satisfying the curiosity of other reviewers as well.)

Table 5-19 in the online Handbook shows that WTP to reduce a microrisk in the chance of sudden death in the current period declines monotonically from $7.91 to $6.04 across the five choice occasions, whereas $7.40 is the overall average WTP estimate in a model that does not control for deviations from the designated choice number or from the median choice time.

We note that the log-likelihood function is maximized when the preference parameters are normalized on the third choice occasion. This may reflect the countervailing influences of fatigue and the development of choice heuristics. Modest order effects have been demonstrated in other similar settings (Bateman et al. (2004)). However, because of the randomized illness assignments in our study, this would not influence our average results.

3) A potential strength of this study is that it should be possible to value both mortality and morbidity effects, and in addition relate these to the different diseases and conditions. But as far as I can see, no such results are given.

RESPONSE: We actually *did* provide numerous examples of morbidity results in the main tables and figures in the original manuscript (i.e. rows 2 through 5 of Table 2 and Table 3, and columns 2 through 5 of Table 4. Clearly we need to discuss this capability of our model in more detail in the text since it was missed by this reviewer. The material at the bottom of page 27 and the top of page 28 should make this capability clearer.

To be specific, Tables 2 and 3 in the paper show results for five different illness profiles with no latency: sudden death scenarios as well as for "one year sick, non-fatal," "five years sick, nonfatal," "one year sick, then die," and "five years sick, then die." Table 4 is even more general, covering the same five illness profiles, but for 35-year-olds and for 65-year-olds, with different amounts of *latency* before the morbidity period commences. We now emphasize and discuss our morbidity results in several new places within the manuscript: at the bottom of page 29 and top of page 30, where we talk about illness profiles 2 through 5; Section (iv) on page 32 (WTP as a function of disease

latency); Section (vi) on page 35 (WTP to reduce risk of other illness profiles as a function of age).

4) Mortality effects could, it appears, be differentiated by cause of death etc; which seems not to have been done. Neither are any valuation results, related directly to reduced morbidity, reported. (On page 12 of the paper it is however stated that more discussion of such issues is found in "other work". This is understandable given the magnitude of the effort behind the material presented here. If there is a conscious decision to shy away from such issues here, it should however be stated more clearly and upfront; from my perspective the current version otherwise appears as quite incomplete.)

RESPONSE: Again, it is incorrect that "Neither are any valuation results, related directly to reduced morbidity, reported." See our response to the last comment.

In addition to the analysis that is here, an extended analysis of the systematic effects of the twelve different illness names (across males and females) seemed beyond the capacity of a single journal article. If *k* is the number of distinct illness labels, then *k-1* illness-specific indicator variables need to be interacted with at least four baseline variables in the model (the "intercept" and the sick-years, recovered-years, and lost life-years variables). The parameter space expands rapidly. A separate paper to discuss the effect of specific illnesses on WTP for micro-risk reduction is currently in the review process and available from the authors. For the reader, we now discuss this in the introduction (on page 4: "…we leave to related and future papers a more-detailed exploration of the roles played by, for example, age, current health status, specific-illness effects, subjective risk beliefs, choice set complexity and alternative discounting assumptions.") and in footnote 51 on page 27.

Our "valuation results, related directly to reduced morbidity" were reported in the original manuscript, as noted above, but we now make an effort to draw more attention to these results in the tables and in the figures.

5) Discounting is an important aspect of the study, in particular since many of the attributes to be valued refer to effects occurring in the future (often distant). Since valuations at both a short and long time range are simultaneously derived, implicit discount rates from these choices can, in principle, be derived endogenously. Surprisingly, no effort to derive implicit discount rates seems to have been done. Instead, alternative assumptions regarding discounting are adopted by the authors: 3, 5 and 7 percent, which are taken to be constant over time. I fail to understand how such assumptions can be consistently made, when individuals' choices at the same time must need to imply particular rates of discounting. Also, the assumed constancy of discount rates accords poorly with observed patterns from related studies (see e g Cropper et al (1994)). I think the richness of results from this survey could, much more actively and constructively, be utilized in this context.

RESPONSE: We agree that discounting is extremely important. We actually devoted a great deal of attention to discounting at several stages of this research project but simply could not squeeze a separate discussion of this topic into this paper. One important thing to keep in mind is that our sample involves a single cross-section of respondents. It does

not follow the same individuals over time so that we might infer how their discounting behavior changes as they age.

For the Knowledge Networks panel, we have actually calculated individual-specific exponential financial discount rates based on a separate survey of the same population. For that other sample, we estimated a model to explain individual discount rates in terms of individual attributes (including age effects that vary with gender, education, and income, separate gender , education and income effects, and subjective life expectancy. We have explored models which transfer this fitted model for individual discount rates to the different sample of KN panelists used for this survey. Table 5-11 in the online Handbook compares our results assuming a "5% discount rate for all" to those when we impose the fitted individual discount rates from our other KN survey.

These models, however, assume all of the discounting information is deterministic, which is troublesome. We have another project in the works where we specify a joint model to explain both these health-related choices and choices among time-denominated receipt of money by pooling our data across multiple surveys of the same population (on different topics). However, this is a daunting jointly estimated maximum likelihood problem, given the uniqueness of everyone's illness profiles. So far, the algorithm has been balky, although we hope to return to that model in the future.

We discuss our discounting efforts now in footnote 30 on page 15 and footnote 38 on page 20. We also discuss the effects of differing discount rate assumptions in Section B (ii) starting on page 30. Also see Section 5.6.5 of the online Handbook. In brief, Table 5-12 in the online Handbook shows the effects on all of our simulated WTP estimates for different illness profiles in the main paper as we change the discounting assumptions. The first column of results is for the individual-specific discount rates. The next three columns show the results for the 3%, 5%, and 7% assumptions. In the final column, however, we show what happens if we estimate the model based on the fitted individual discount rates and then simulate WTP under the counterfactual conditions where everyone has a 5% discount rate. This last set of results corresponds very closely to the results when a 5% discount rate overall is imposed.

To the extent that we subscribe to "paternalistic libertarian" policy-making, we may want to override some of the higher subjective financial discount rates that may be present for our respondents. However, it would indeed be much better to have individual-specific information about discount rates, rather than to have to transfer a model from another sample. We have opted for full disclosure of the consequences of these different discounting assumptions in our online Handbook. We need to emphasize, however, that our investigations concerning discount rates still seem to be far more comprehensive than any other research on mortality- and morbidity-related benefits estimation to date.

6) I cannot see that a proper scope test is performed in this study (I may be wrong). If not it is a major weakness. In my view, a positive outcome of a scope test should not simply be an indication that WTP increases with the good to be valued, but (here) that value is roughly proportional to scope (at least, for small probabilities and/or life extensions).

RESPONSE: In fact, this study may embody the ultimate in scope tests. Every survey instrument is essentially unique in its set of attribute levels. The risks are varied randomly

across people within each named illness and across named illnesses within each individual. All other attributes are also varied randomly both within and between individuals. It is hard to imagine any more scope-test-ready survey. The standard external scope test requires only that programs of different magnitudes be offered to different people. For example, we might have offered one risk reduction to one-half of the sample and a different risk-reduction to the other half of the sample, and then sought to determine whether WTP was larger for the larger risk reduction. Instead, we use a variety of different sizes of risk reductions and describe them relative to a variety of different baseline risks, and we also randomly vary the characteristics of each risk reduction. We demonstrate ample sensitivity of WTP to a large number of program attributes that vary across individuals.

As expected, we find that WTP is roughly proportional to scope, even if we define "scope" narrowly as just the size of the risk reduction, without the further introduction of significant heterogeneity in types of risks. Despite the wide variety in original and reduced absolute levels of risk, there are only six values for the actual risk *difference* in this study. By construction (as an artifact of our structural utility-maximization model and the discounted expected utility-differences upon which choices are based) WTP is approximately proportional to the size of the risk change. The only disparity stems from the form assumed for the *yterm* and *cterm* portions of our model, which reflect the assumed future time profiles of income and program cost according to whether the individual gets sick from the specified illness or stays healthy. For the reader, we mention the scope test in the middle of page 12 and in the introduction to Section III on page 24. Also see Section 2.5 of the online Handbook.

7) More questioning should also be done of the basic premise, to use a direct monetary payment ("co-pay") as the payment vehicle. In my view the authors take too lightly on the issue of accepting such a condition; it is done by simply referring to their claim that the health goods entering into the survey are not covered by public health or insurance schemes. In my view this is too weak: there is no compelling reason why individuals should necessarily adopt such a presumption if it does not accord with their own experiences and reality.

RESPONSE: We asked our respondents to assume that "Your participation in a program would cost money. These higher costs might take the form of a co-payment when you visit your doctor visit or higher monthly health insurance costs." (Form 21)

We selected this payment vehicle based on our dozens of one-on-one think-aloud trial runs with test subjects (on different occasions throughout the survey development phase) at the Knowledge Networks facility in Menlo Park, CA, and our extensive debriefings of these randomly selected subjects. People were familiar and comfortable with this payment vehicle, based on their own experiences. For younger patients, many diagnostic tests require co-pays because they are not fully covered by medical insurance. Even for older patients, many of the newer or less common diagnostic tests must be paid for out-of pocket. More generally, nearly all private medical insurance requires a visit co-pay just to see a doctor. Based on our focus groups, pretesting and early scenario rejection analysis, the co-payment approach was both acceptable and familiar to respondents.

Is it also reasonable to think that individuals will believe that the required payment will be charged, throughout the person's future life span (as is assumed here)? I think, maybe not. If so, payments as expressed in the survey will be, in reality, lower.

> RESPONSE: It is not the case that we "…think that individuals will believe that the required payment will be charged, throughout the person's future life span (as is assumed here)?" Instead, we impute that respondents do not expect to pay the annual cost while they are sick or if they die early from the specified health threat. At the bottom of page 15, we state that "…individuals are assumed to anticipate paying ("*p*") program costs only when they are neither sick nor dead,…". This assumption accounts in part for the complexity of the *cterm* and *yterm* parts of model. It is certainly true that under the alternative and less plausible assumption that this referee thinks we made, the estimating specification would have been simpler and the WTP calculations would have been easier.
>
> We agree, however, that respondent inferences about payment trajectories under different conditions are a reasonable concern and we now evaluate empirically the impact of alternative assumptions. If the cost of the program is a copayment for a privately provided medical service, the respondent would certainly not need to pay once they had the illness *unless* they were being tested for a risk of recurrence. We now evaluate and report the empirical consequences of altering this assumption so that readers are both aware of this issue that you raise and understand its empirical implications. Also see page 16 ("We also assume that the individual does not expect to pay the cost of the program if they are currently experiencing that illness or if they die from the illness.") and footnote 32 for additional discussion.
>
> We now also explore the implications of private versus public risk reduction programs. For public programs, it is likely that the individual will have to continue to bear the costs even if they suffer the illness. Table 5-17 in the online Handbook demonstrates that WTP is actually not much affected by whether the respondent is assumed to bear the program's costs while sick. Many illnesses are of short duration relative to nominal life expectancy in the absence of the illness, and WTP for an illness includes the disutility of recovered time (if any) and lost life-years, which seem to dominate the story in many cases.

I will now discuss the usefulness of the results derived in the paper, for public policy in areas involving mortality and morbidity risks. This is indeed one of the main touted objectives of this exercise. I agree that the results of this survey can (after further refinement), hopefully, be put to good use in attempting to answer the following question: What is the value of additional lifetime years (possibly, expressed as a VSL measure) to individuals, when these individuals, as valued by the individuals themselves? This study adds (I believe, in a productive way) to this literature.

> RESPONSE: As you point out, we do estimate the value of future lost life year under a wide range of assumptions. We choose to emphasize that particular area of overlap between our approach and the existing literature because this overlap is where we can "validate" our model and its resulting WTP estimates against prior work. But we hasten to point out that the estimation framework in this paper is far more general than just this. We derive WTP measures for a wide variety of sequences of health states, including the 45 different illness profiles in Table 4, with either one or five years of morbidity, five

different latencies for 35-year-olds, and three different latencies for 65-year-olds. We also illustrate age patterns for three potentially policy-relevant illness profiles in addition to "sudden death now" in Figures 2, 3, and 5. Estimating the "VSL" (for purposes of cross-validation) is just one special case of the range of scenarios our model can accommodate. Our model can answer many more questions than just "What is the value of additional lifetime years?"

This is however different from the following question: How does *society* value the loss of the same potential lifetime years? This is the relevant question in a public policy context. For a number of reasons there may be (and likely are) discrepancies between these individualistic and social values, stemming from (paternalistic and purely selfish) inter-personal preferences, and from financial and economic externality effects; some of these are discussed by Strand (2006), but there are several others as well. It would be unfair to the authors of this paper to demand that a comprehensive discussion of such further issues be included. My point here is more mundane, namely to point out the limitations of the study in this regard.

RESPONSE: Yes, we agree that we only estimate individuals' self-regarding preferences in this paper. We do this because, presently and historically, EPA has utilized primarily self-regarding preferences in its valuation of the health risk reductions. The present study, like many others, is based upon the notion of consumer sovereignty—the amount people are willing to pay to reduce their own risks of experiencing specific patterns of adverse future health states.

To your point, we also agree that individuals' self-regarding and other-regarding preferences for risk reductions may diverge. In our other work with our "public choices" survey, we find depressingly little evidence of willingness to pay for medical treatments for others, unless it is for an illness in which the respondent him/herself has some self-interest. But that is a different paper based on a fundamentally different survey instrument administered to a separate sample. It is currently under review at the *Journal of Health Economics*.

This present paper is intended as the "flagship" of the suite of papers from our private-choices study, and all other papers refer to it. We can concede that it does not solve every challenge that policy-makers might ever face in the task of measuring the social benefits of risk reductions, but we are confident that we have taken a large step forward in this process. We hope that future researchers will take our concept and framework and improve upon it, given what we have learned and given the copious documentation of our study that is now contained in the new online Handbook.

I think it is also a bit overblown when the authors claim that their study opens up a new main perspective on VSL being, in reality, differentiated across age groups etc. Many studies do this. One must not confuse the debate going on in stated preference research on VSL, with the (rough and crude) application of VSL measure in public policy contexts (where a single number must usually be chosen; and where I suspect that at least informed parties are fully aware of the limitations on such a principle, and the relationships considered here).

RESPONSE: In the revised manuscript, we have tried to clarify better our central contribution. It is not merely to explore sources of heterogeneity across individuals (i.e.,

by age group). Rather, it is to specify and estimate individuals' WTP for avoided future sick years, recovered/remission years, and lost life years in a way that enables us to value avoiding probabilistic future illness profiles of *any* mix and duration. We do this in the context of a rigorously utility-theoretic structural model of preferences defined over present discounted time periods in *different* health states over the respondent's remaining life, rather than in a reduced-form specification. What is innovative is the more-general, but still utility-theoretic, structure of our model. Our age effects, for example, involve heterogeneity not just in a reduced form measure of WTP, but in several (although not all) of the different marginal utility-related parameters that make up our translog-type flexible local approximation to preferences.

A construct that approximates "the VSL" is just one special case of the continuous spectrum of illness profiles that can be accommodated by our model.

It is clear that you have given this paper an exceeding close read, so our failure to make clear our innovations has prompted us to rewrite parts of the introduction.

I also have a few smaller, and editorial, comments.

In eq. 1, page 16, a very simple linear version of an individual indirect utility function is presented. While modified somewhat later, I think this gives a wrong impression: in reality all components of this function exhibit dependence. In addition, the value of income will depend on state. Some of these issues are addressed in appendix E, but, I suspect, not all. The attempt to make the survey as general as possible here, I think, clashes with an objective to find a best possible utility function that can be, practically, estimated.

RESPONSE: We agree that utility as a function of different sequences of future health states is a fascinating economic question. We start with a simple and straightforward linear and additively separable model as a way to introduce the basic mechanics of discounted expected utility in this choice context. Once that hurdle is cleared, however, we have to acknowledge that this initial simple specification is clearly rejected by the data, so we generalize the specification to accommodate features such as diminishing marginal utility and the dependence of some marginal utilities on the levels of other variables (the interaction terms).

We are not sure that we see a "clash" between collecting a rich array of information that is important to people's choices among risk-reduction programs, and specifying a model that makes good use of these data (i.e. by relaxing restrictive assumptions that are clearly rejected by the data). The complexity in our specification is not gratuitous or optional; it is necessary and warranted by the data. Simple models are of course desirable when they are not rejected by the data, and we are certainly aware that this research would have been published much more quickly if it was simpler. Unfortunately, any simpler model is wrong. If we had used a simpler survey that had elicited less information from our respondents, of course, we might have been oblivious to this fact and could blissfully limit our analyses to simpler specifications.

With respect to the potential dependence of the marginal utility of discounted future net income on discounted future health states, we have explored a wide variety of specifications that allow the marginal utility of net income to depend upon the pattern of future health states associated with the illness profile in question. Only in the crudest

models is there statistically significant dependence of the marginal utility of income on the health states in the illness profile: if we allow the coefficient that captures the marginal utility of transformed net income to vary with the *three* basic discounted health-state terms in the model, *pdvi*, *pdvr*, and *pdvl*, only the coefficient on the interaction with the discounted lost life-years term is significant, and only at the 10% level. When we expand the model to include interactions that allow the marginal utility of transformed net income to vary with all of the health-state-related terms in the model (i.e. base terms, squared terms, and interactions), only one of the coefficients achieves an asymptotic absolute t-test value of greater than 1.27. Furthermore, the maximized value of the log-likelihood increases by only five points for these additional 11 parameters. The chi-squared test statistic is only 9.786, whereas even the 10% critical value is 17.275.

One would need to elicit a lot more information about people's specific expectations regarding their future income trajectories, both in a continued healthy state and under each proposed illness scenario, before it would be possible to get to the bottom of the question how people's future marginal utilities of income will depend upon their future health status, conditional on each illness profile. The survey duration covered by our agreement with Knowledge Networks did not permit us to ask every question that would have provided information we could use.

It might be possible to contemplate a model that goes back to our starting point of utility in each future period, before we make the shift to a model where individuals view the *discounted* health profiles as the objects of choice in the current-period decision about signing up to participate in an annual diagnostic program for the indefinite future. We do this sort of thing in another (working) paper called "Two Types of Age Effects." The marginal utility of income in each future period could be modeled as depending upon health status in each future period, before discounting back to the present. However, that is a substantially more complicated model, and one would shudder at the thought of having to untangle the fitted utility specification to solve for WTP. However, this might go on our agenda of other more-general models to consider in the future.

On page 16 also, individuals are stated to be informed about baseline risks. Do they believe in these statements? In particular, is there debriefing data illuminating this?

RESPONSE: Many respondents did accept the baseline risks. To evaluate this, and control for those who did not, we posed follow-up questions after each choice scenario that allowed respondents to indicate when they thought each program might begin to benefit them, or whether they expected never to benefit from the program. We also asked preliminary questions about subjective risks related to each specific illness. In the current paper, however, we lean upon the randomization of risks and illness profiles across individuals to preclude omitted variables bias in our estimates.

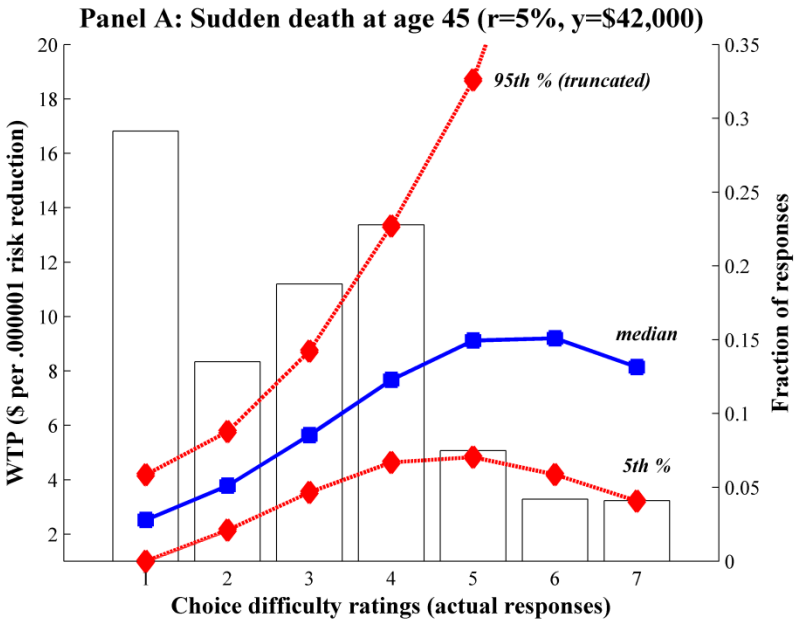
In other work, "The Effect of Health Status on Willingness to Pay for Morbidity and Mortality Risk Reductions," we explore the systematic effects on WTP of same-illness and other-illness subjective risks and same-illness and other-illness current morbidity. Again, there is enough material in this question to fill a 61-page manuscript. Our randomized design at least ensures that the estimates in the current paper reflect the central tendency across the *distribution* of beliefs about risks and illness profiles in the current paper.

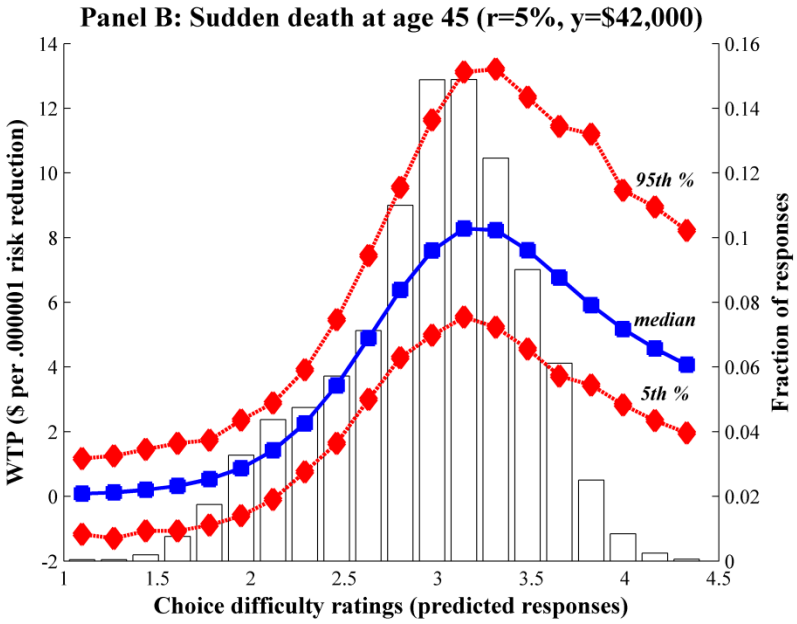
One exception, however, is our new correction for those cases wherein individuals indicated that they would "never" benefit from a particular program. This is one of our new suite of "scenario adjustment/rejection" corrections. For each risk-reduction program, we have explored interaction terms between all of the main variables in the model and this "would never benefit" indicator. In our simulations, we set this indicator to zero, implying conditions where nobody believed the program would never benefit them. Our WTP estimates are thus purged of any systematic bias from this type of scenario rejection.

In other work ("Scenario Adjustment in Stated Preference Surveys," *Journal of Choice Modeling*), we advocate strongly for the importance of debriefing questions when respondents are asked to react to hypothetical choice situations. In that paper, we also explored the use of a measure of the "minimum overestimate of the latency" of the illness in question, constructed from the debriefing question about when benefits would start. However, the grounds for such an adjustment are probably a bit shakier, because of the possibility of ambiguity in this follow-up question. Perhaps our question should have been more direct. In the current paper, we found that including adjustments for the "minimum overestimate of the latency" produced implausible simulated WTP amounts for our end-of-life effects summarized in Figure 3. In this case, the adjustments produced many negative WTP amounts, suggesting that the adjustment may be too aggressive. Thus we have backed off from using this adjustment, although we now use other appropriate adjustments, including two that replace our earlier strategy of disqualifying some observations.

The survey contains a debriefing part, but little information is provided about the debriefing process, what it contains, and the results from it.

RESPONSE: As noted above, we now explicitly include one of the debriefing questions after each choice set as an integral correction in our new specification. Some of the debriefing questions from the rest of the survey form the basis of separate papers. For example, "Subjective choice difficulty in stated choice tasks" is currently in the review process. In that paper, we explore how WTP varies systematically with actual and fitted choice difficulty ratings. Below are two figures from that paper, although bear in mind that the utility specification for the model which yielded these results is a previous-generation model without the same scenario adjustment/rejection controls employed here. Thus the results depicted in these figures must be considered to be preliminary.





These figures illustrate that the modal *actual* choice-difficulty rating was 1 = least difficult. As rated difficulty increases, so does fitted WTP, at least up until a point, after which WTP tends to decline.

In the current paper, of course, the WTP amounts do indeed reflect the *distribution* across the sample in choice difficulty ratings, and we rely on the extensive randomization of illnesses and illness profiles across each all of the choice sets in the survey to eliminate systematic biases in our estimates.

As for editing and presentation, I think the paper makes for a heavy read, as the authors try to present almost too much material (a variety of different sets of variables included for example).

A reformulation should focus more on the essentials: one tries to value both VSL and morbidity, and for different reasons for these. Thus also some sort of QALY concept is involved here, and a specific link to the discussion of the latter concept would then preferably have been in order.

> RESPONSE: We have tried to streamline the paper in several ways. Table 1 involves only two specifications. Model 1 is rudimentary, to prove that our statistically significant coefficients are not merely an artifact of specification searching, and Model 2 is the specification we use to generate our WTP estimates. If, by "different sets of variables" this reviewer is referring to the interaction terms and the age heterogeneity warranted by the data, it is clear that exclusion of these variables is strongly rejected. In the original paper, the nine additional terms improved the log likelihood by more than 34. There are strong arguments in the existing literature for allowing quadratic age effects, and a strong a priori argument for allowing the effect of lost life-years to vary systematically with the extent of prior morbidity (e.g. allowing for "fates worse than death").
>
> Our current paper does value both lost life years and morbidity. We now briefly mention QALY measures on page 27. In current work, we looking more closely at the time trade-off between sick-years and lost life-years that has been popular in the QALY literature. We are working with a current Ph.D. student for whom this task will constitute a dissertation chapter. This is certainly a big enough question to motivate an entirely new paper. The current manuscript anticipates this next enterprise.

There should be a more explicit statement, upfront, concerning any possible sharing of material with other papers/publications coming out of this project. This may also make it possible to concentrate attention on those aspects that are the center of attention in this presentation.

> RESPONSE: This is indeed the flagship paper that is meant to be the centerpiece from which all of our other papers are derived. We have rewritten this portion of the introduction (see the end of the paragraph at the top of page 4).

Also, it will probably not be feasible to include all appendices A-F in any final published version of the paper. Possibly, material from appendix E on model estimation ought to be included in a possible published version in JEEM, but not the other appendices. This would then, clearly, in case need to be reflected in the presentation.

> RESPONSE: We have been asked by another reviewer to combine the appendices in a working paper which will be published online along with the paper. We hope this comports with your request as well. The perception that it is "probably not feasible" to include all of the appendices in the final published version is accurate. The detailed explanations and auxiliary specifications we have generated in response to these current reviewers and other previous reviewers now amount to a book-sized manuscript.

8 INDEX

9 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., 2003. Age variation in workers' value of statistical life.

Ashenfelter, O., Greenstone, M., 2004. Estimating the value of a statistical life: The importance of omitted variables and publication bias. *American Economic Review* 94, 454-460.

Bandilla, W., Bosnjak, M., Altdorfer, P., 2003. Survey administration effects? A comparison of web-based and traditional written self-administered surveys using the ISSP environment module. *Social Science Computer Review* 21, 235-243.

Baron, J., Ubel, P.A., 2002. Types of inconsistency in health-state utility judgments. *Organizational Behavior and Human Decision Processes* 89, 1100-1118.

Bateman, I.J., Cole, M., Cooper, P., Georgiou, S., Hadley, D., Poe, G.L., 2004. On visible choice sets and scope sensitivity. *Journal of Environmental Economics and Management* 47, 71-93.

Berrens, R.P., Bohara, A.K., Jenkins-Smith, H., Silva, C., Weimer, D.L., 2003. The advent of Internet surveys for political research: A comparison of telephone and Internet samples. *Political Analysis* 11, 1-22.

Best, S.J., Krueger, B., 2002. New approaches to assessing opinion: The prospects for electronic mail surveys. *International Journal of Public Opinion Research* 14, 73-92.

Best, S.J., Krueger, B., Hubbard, C., Smith, A., 2001. An assessment of the generalizability of internet surveys. *Social Science Computer Review* 19, 131-145.

Birnbaum, M.H., 2004. Human research and data collection via the Internet. *Annual Review of Psychology* 55, 803-832.

Breslow, N.E., Day, N.E., 1980. *Statistical methods in cancer research: the analysis of case-control studies*. IARC, Lyon.

Cameron, T.A., 2010. Euthanizing the Value of a Statistical Life. *Review of Environmental Economics and Policy* 4, 161-178.

Cameron, T.A., Crawford, G.D., 2003. Independent dimensions of sociodemographic variability in neighborhood characteristics at the tract level of the 2000 Census, Department of Economics, University of Oregon, Eugene, OR.

Cameron, T.A., DeShazo, J.R., 2010a. Comprehensive (as opposed to end-stage) sample selection in a major internet consumer panel, Department of Economics, University of Oregon, Eugene, OR.

Cameron, T.A., DeShazo, J.R., 2010b. Differential Attention to Attributes in Utility-Theoretic Choice Models. *Journal of Choice Modelling* 3, 73-115.

Cameron, T.A., DeShazo, J.R., 2011. Demand for Health Risk Reductions, Working Papers, Department of Economics, University of Oregon, Eugene, OR 97403.

Cameron, T.A., DeShazo, J.R., Johnson, E.H., 2010a. The effect of children on adult demands for health-risk reductions. *Journal of Health Economics* 29, 364-376.

Cameron, T.A., DeShazo, J.R., Johnson, E.H., 2010b. Scenario adjustment in stated preference research. *Journal of Choice Modelling* 4, 9-43.

Cameron, T.A., DeShazo, J.R., Johnson, E.H., 2011. Willingness to Pay for Health Risk Reductions: Differences by Type of Illness, Working paper, Department of Economics, University of Oregon, Eugene, OR 97403.

Cameron, T.A., DeShazo, J.R., Stiffler, P., 2010c. Demand for health risk reductions: A cross-national comparison between the U.S. and Canada. *Journal of Risk and Uncertainty* 41, 245-273.

Carson, R.T., Groves, T., List, J., Machina, M., 2004. Probabilistic influence and supplemental benefits: a field test of the two key assumptions underlying stated preferences, Department of Economics, University of California, San Diego, CA.

Chamberlain, G., 1980. Analysis of covariance with qualitative data. *Review of Economic Studies* 47, 225-238.

Chestnut, L.G., Rowe, R.D., Lazo, J.K., Breffle, W.S., 2003. Economic valuation of mortality risk reduction: Stated preference approach. Prepared for Cornell University by Stratus Consulting Inc., Boulder, CO.

Collett, D., 2003. *Modelling binary data*, 2nd ed. Chapman & Hall/CRC, London.

Corso, P.S., Hammitt, J.K., Graham, J.D., 2001. Valuing mortality-risk reduction: Using visual aids to improve the validity of contingent valuation. *Journal of Risk and Uncertainty* 23, 165-184.

Cuddeback, G., Wilson, E., Orme, J.G., Combs-Orme, T., 2004. Detecting and statistically correcting sample selection bias. *Journal of Social Service Research* 30, 19-33.

Cummings, R.G., Taylor, L.O., 1999. Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method. *American Economic Review* 89, 649-665.

Das, M., Newey, W.K., Vella, F., 2003. Nonparametric estimation of sample selection models. *Review of Economic Studies* 70, 33-58.

de Blaeij, A., Florax, R., Rietveld, P., Verhoef, E., 2003. The value of statistical life in road safety: a meta-analysis. *Accident Analysis and Prevention* 35, 973-986.

de Bruin, W.B., Keren, G., 2003. Order effects in sequentially judged options due to the direction of comparison. *Organizational Behavior and Human Decision Processes* 92, 91-101.

DeShazo, J.R., Cameron, T.A., 2006a. The Effect of Health Status on Willingness to Pay for Morbidity and Mortality Risk Reductions, Working Paper, Department of Economics, University of Oregon, Eugene, OR 97403.

DeShazo, J.R., Cameron, T.A., 2006b. Two Types of Age Effects in the Demand for Reductions in Mortality Risks with Differing Latencies, Working Paper, Department of Economics, University of Oregon, Eugene, OR 97403.

Deutskens, E., de Jong, A., de Ruyter, K., Wetzels, M., 2006. Comparing the generalizability of online and mail surveys in cross-national service quality research. *Marketing Letters* 17, 119-136.

Dow, W.H., Philipson, T.J., Sala-I-Martin, X., 1999. Longevity complementarities under competing risks. *American Economic Review* 89, 1358-1371.

Duffy, B., Smith, K., Terhanian, G., Bremer, J., 2005. Comparing data from online and face-to-face surveys. *International Journal of Market Research* 47, 615-639.

Duquette, E., Cameron, T.A., DeShazo, J.R., 2010. Subjective choice difficulty in stated choice tasks, Working Paper, Department of Economics, University of Oregon, Eugene, OR 97403.

Eeckhoudt, L.R., Hammitt, J.K., 2004. Does risk aversion increase the value of mortality risk? *Journal of Environmental Economics and Management* 47, 13-29.

Ehrlich, I., 2000. Uncertain lifetime, life protection, and the value of life saving. *Journal of Health Economics* 19, 341-367.

Ehrlich, I., Chuma, H., 1990. A model of the demand for longevity and the value of life extension. *Journal of Political Economy* 98, 761-782.

Evans, J.R., Mathur, A., 2005. The value of online surveys. *Internet Research-Electronic Networking Applications and Policy* 15, 195-219.

Eysenbach, G., 2004. Improving the quality of web surveys: The checklist for reporting results of Internet e-surveys (CHERRIES). *Journal of Medical Internet Research* 6, 12-16.

Fetherstonhaugh, D., Slovic, P., Johnson, S.M., Friedrich, J., 1997. Insensitivity to the value of human life: A study of psychophysical numbing. *Journal of Risk and Uncertainty* 14, 283-300.

Fitzgerald, J., Gottschalk, P., Moffitt, R., 1998a. An analysis of sample attrition in panel data - The Michigan Panel Study of Income Dynamics. *Journal of Human Resources* 33, 251-299.

Fitzgerald, J., Gottschalk, P., Moffitt, R., 1998b. An analysis of the impact of sample attrition on the second generation of respondents in the Michigan Panel Study of Income Dynamics. *Journal of Human Resources* 33, 300-344.

Fox, C.R., Irwin, J.R., 1998. The role of context in the communication of uncertain beliefs. *Basic and Applied Social Psychology* 20, 57-70.

Goritz, A.S., 2004. Recruitment for online access panels. *International Journal of Market Research* 46, 411-425.

Graham, D.A., 1981. COST-BENEFIT-ANALYSIS UNDER UNCERTAINTY. *American Economic Review* 71, 715-725.

Greene, W.H., 2008. *Econometric Analysis*, 6th ed. Prentice Hall.

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

Hamerle, A., Ronning, G., 1995. Panel analysis for qualitative variables, in: Arminger, G., Clogg, C.C., Sobel, M.E. (Eds.), *Handbook of statistical modeling for the social and behavioral sciences*. Plenum, New York, pp. 401-451.

Hamermesh, D.S., 1985. Expectations, life expectancy, and economic behavior. *Quarterly Journal of Economics* 100, 389-408.

Hammitt, J.K., Graham, J.D., 1999. Willingness to pay for health protection: Inadequate sensitivity to probability? *Journal of Risk and Uncertainty* 18, 33-62.

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.

Harrison, G.W., Rutstrom, E.E., 2006. Eliciting subjective beliefs about mortality risk orderings. *Environmental & Resource Economics* 33, 325-346.

Hausman, J.A., Wise, D.A., 1979. Attrition bias in experimental and panel data - Gary income-maintenance experiment. *Econometrica* 47, 455-473.

Heckman, J.J., 1979. Sample selection bias as a specification error. *Econometrica* 47, 153-161.

Hosmer, D.W., Jr., Lemesow, S., 2000. *Applied logistic regression*, 2nd ed. Wiley, New York.

Ilieva, J., Baron, S., Healey, N.M., 2002. Online surveys in marketing research: pros and cons. *International Journal of Market Research* 44, 361-376.

Johnson, F.R., Banzhaf, M.R., Desvousges, W.H., 2000. Willingness to pay for improved respiratory and cardiovascular health: A multiple-format, stated-preference approach. *Health Economics* 9, 295-317.

Krupnick, A., Alberini, A., Cropper, M., Simon, N., O'Brien, B., Goeree, R., Heintzelman, M., 2002. Age, health and the willingness to pay for mortality risk reductions: A contingent valuation survey of Ontario residents. *Journal of Risk and Uncertainty* 24, 161-186.

Lee, S., 2006. An evaluation of nonresponse and coverage errors in a prerecruited probability web panel survey. *Social Science Computer Review* 24, 460-475.

Leip, D., 2003. Atlas of U.S. Presidential elections. http://www.uselectionatlas.org/.

Lillard, L.A., Panis, C.W.A., 1998. Panel attrition from the panel study of income dynamics - Household income, marital status, and mortality. *Journal of Human Resources* 33, 437-457.

List, J.A., 2001. Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards. *American Economic Review* 91, 1498-1507.

Moxey, A., O'Connell, D., McGettigan, P., Henry, D., 2003. Describing treatment effects to patients: How they are expressed makes a difference. *Journal of General Internal Medicine* 18, 948-959.

Mrozek, J.R., Taylor, L.O., 2002. What determines the value of life? A meta-analysis. *Journal of Policy Analysis and Management* 21, 253-270.

Nevo, A., 2003. Using weights to adjust for sample selection when auxiliary information is available. *Journal of Business & Economic Statistics* 21, 43-52.

Nicoletti, C., Peracchi, F., 2005. Survey response and survey characteristics: microlevel evidence from the European Community Household Panel. *Journal of the Royal Statistical Society Series a-Statistics in Society* 168, 763-781.

O'Conor, R.M., Blomquist, G.C., 1997. Measurement of consumer-patient preferences using a hybrid contingent valuation method. *Journal of Health Economics* 16, 667-683.

Pagan, A., 1984. Econometric issues in the analysis of regressions with generated regressors. *International Economic Review* 25, 221-247.

Pagan, A.R., Nicholls, D.F., 1984. Estimating predictions, prediction errors and their standard deviations using constructed variables. *Journal of Econometrics* 24, 293-310.

Picone, G., Sloan, F., Taylor, D., Jr., 2004. Effects of risk and time preference and expected longevity on demand for medical tests. *Journal of Risk and Uncertainty* 28, 39-53.

Quiggin, J., 2002. Risk and self-protection: A state-contingent view. *Journal of Risk and Uncertainty* 25, 133-145.

Read, D., Loewenstein, G., Rabin, M., 1999. Choice bracketing. *Journal of Risk and Uncertainty* 19, 171-197.

Ridder, G., 1990. Attrition in multi-wave panel data, in: Hartog, J., Ridder, G., Theeuwes, J. (Eds.), *Panel Data and Labor Market Studies*. North-Holland, Amsterdam.

Schillewaert, N., Meulemeester, P., 2005. Comparing response distributions of off line and online data collection methods. *International Journal of Market Research* 47, 163-178.

Schonlau, M., 2004. Will web surveys ever become part of mainstream research? *Journal of Medical Internet Research* 6, 10-11.

Schonlau, M., Zapert, K., Simon, L.P., Sanstad, K.H., Marcus, S.M., Adams, J., Spranca, M., Kan, H.J., Turner, R., Berry, S.H., 2004. A comparison between responses from a propensity-weighted web survey and an identical RDD survey. *Social Science Computer Review* 22, 128-138.

Shogren, J.F., Crocker, T.D., 1999. Risk and its consequences. *Journal of Environmental Economics and Management* 37, 44-51.

Sloan, F.A., Viscusi, W.K., Chesson, H.W., Conover, C.J., Whetten-Goldstein, K., 1998. Alternative approaches to valuing intangible health losses: the evidence for multiple sclerosis. *Journal of Health Economics* 17, 475-497.

Smith, T.W., 2003. An experimental comparison of knowledge networks and the GSS. *International Journal of Public Opinion Research* 15, 167-179.

Smith, V.K., Evans, M.F., Kim, H., Taylor, D.H., 2004. Do the near-elderly value mortality risks differently? *Review of Economics and Statistics* 86, 423-429.

Taylor, L.O., Morrison, M.D., Boyle, K.J., 2010. Exchange rules and the incentive compatibility of choice experiments. *Environmental & Resource Economics* 47, 197-220.

Tourangeau, R., 2004. Survey research and societal change. *Annual Review of Psychology* 55, 775-801.

Ubel, P.A., Richardson, J., Baron, J., 2002. Exploring the role of order effects in person trade-off elicitations. *Health Policy* 61, 189-199.

Vella, F., 1998. Estimating models with sample selection bias: A survey. *Journal of Human Resources* 33, 127-169.

Viscusi, W.K., 1993. The value of risk to life and health. *Journal of Economic Literature* 31, 1912-1946.

Viscusi, W.K., 2003. Racial differences in labor market values of a statistical life. *Journal of Risk and Uncertainty* 27, 239-256.

Viscusi, W.K., Aldy, J.E., 2003. The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty* 27, 5-76.

Winship, C., Mare, R.D., 1992. Models for sample selection bias. *Annual Review of Sociology* 18, 327-350.

Wright, K.B., 2005. Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and Web survey services. *Journal of Computer-Mediated Communication* 10.

Yeung, R.Y.T., Smith, R.D., McGhee, S.M., 2003. Willingness to pay and size of health benefit: an integrated model to test for 'sensitivity to scale'. *Health Economics* 12, 791-796.

10 Appendix: One instance of the randomized survey instrument

We append to this document one particular instance of our survey instrument. Keep in mind that each survey instrument was tailored to the gender and current age of the respondent. Gender matters because males may see “prostate cancer” among the ten illness labels drawn randomly from the list of eleven possibilities for males. Females may see “breast cancer.” Age matters because no respondent was allowed to consider an illness profile described as beginning in the past, when they were younger that they were at the time they took the survey. Illnesses needed to be future illnesses.

Our example survey instrument is available only in .pdf form, so it must be appended to this document as a separate step. If the survey instrument is missing from the current copy of this document, please do not hesitate to request that it be supplied separately.

Welcome

We want to learn about how you view threats to your health.

Your answers may help public officials provide you with better ways of managing your health.

Please take your time.

{Form 1 - Private}

Continue

How much does each of the following threaten your health?

Select one answer from each row in the grid

	Very little 1	2	3	4	A great deal 5
Unsafe foods	○	○	○	○	○
Unsafe working conditions	○	○	○	○	○
Violent crime	○	○	○	○	○

	Very little 1	2	3	4	A great deal 5
Unsafe drinking water	○	○	○	○	○
Poor air quality	○	○	○	○	○
Unsafe roads	○	○	○	○	○

	Very little 1	2	3	4	A great deal 5

{Form 2 - Private}

Next Question

Have you, or a family member or friend, suffered from any of the following?

Select all answers that apply in the grid

	I have	Family or friends have
Respiratory disease - (asthma, emphysema, bronchitis)	☐	☐
Diabetes	☐	☐
Alzheimer's disease	☐	☐
Heart Disease -(heart attack, angina)	☐	☐

	I have	Family or friends have
Cancer - (colon, breast, prostate, etc.)	☐	☐
Stroke - (stroke, blood clot, aneurysm)	☐	☐
Major car accident	☐	☐

	I have	Family or friends have

{Form 3 - Private}

Next Question

Have you, or a family member or friend, experienced any of the following?

Select all answers that apply in the grid

	I have	Family or friends have
High cholesterol levels	☐	☐
High blood pressure	☐	☐
Extended hospitalization	☐	☐

	I have	Family or friends have
Major surgery	☐	☐
Periods of moderate to severe pain	☐	☐

	I have	Family or friends have

{Form 4 - Private}

Next Question

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?

Select one answer from each row in the grid

	Low risk 1	2	3	4	High risk 5
Respiratory disease - (asthma, emphysema, bronchitis)	○	○	○	○	○
Diabetes	○	○	○	○	○
Alzheimer's disease	○	○	○	○	○
Heart Disease -(heart attack, angina)	○	○	○	○	○

	Low risk 1	2	3	4	High risk 5
Cancer - (colon, breast, prostate, etc.)	○	○	○	○	○
Stroke - (stroke, blood clot, aneurysm)	○	○	○	○	○
Major car accident	○	○	○	○	○

	Low risk 1	2	3	4	High risk 5

{Form 5 - Private}

Next Question

Is there room for you to reduce your health risks by improving your lifestyle or habits in these ways?

Select one answer from each row in the grid

	No room to improve 1	2	3	4	Much room to improve 5
drink less alcohol	○	○	○	○	○
quit smoking	○	○	○	○	○
eat a healthier diet	○	○	○	○	○
see a doctor more regularly	○	○	○	○	○

	No room to improve 1	2	3	4	Much room to improve 5
exercise more	○	○	○	○	○
lose weight	○	○	○	○	○
use a seat belt more	○	○	○	○	○

	No room to improve 1	2	3	4	Much room to improve 5

{Form 6 - Private}

Next Question

Changing your lifestyle or habits can be difficult because it requires time, money, and effort.

How difficult would it be for you to do the following things?

Select one answer from each row in the grid

	easy to do 1	2	3	4	hard to do 5
drink less alcohol	○	○	○	○	○
quit smoking	○	○	○	○	○
eat a healthier diet	○	○	○	○	○
see a doctor more regularly	○	○	○	○	○

	easy to do 1	2	3	4	hard to do 5
exercise more	○	○	○	○	○
lose weight	○	○	○	○	○
use a seat belt more	○	○	○	○	○

	easy to do 1	2	3	4	hard to do 5

{Form 7 - Private}
{Displays only those rows for which there is some "room to improve" on previous screen}

Next Question

How much do you think that improving your lifestyle or habits would reduce your risk of:

Select one answer from each row in the grid

	Very little 1	2	3	4	A lot 5
Respiratory disease - (asthma, emphysema, bronchitis)	○	○	○	○	○
Diabetes	○	○	○	○	○
Alzheimer's disease	○	○	○	○	○
Heart Disease -(heart attack, angina)	○	○	○	○	○

	Very little 1	2	3	4	A lot 5
Cancer - (colon, breast, prostate, etc.)	○	○	○	○	○
Stroke - (stroke, blood clot, aneurysm)	○	○	○	○	○
Major car accident	○	○	○	○	○

	Very little 1	2	3	4	A lot 5

{Form 8 - Private}

Next Question

Doctors tell us that someone like you, who is now about 54 years old, can expect to live until about 88. (Later, we will ask how long you think you will live.) In this survey we focus on health programs that reduce your risk of getting sick and dying in the 34 years between now and age 88.

{Form 9 - Private}

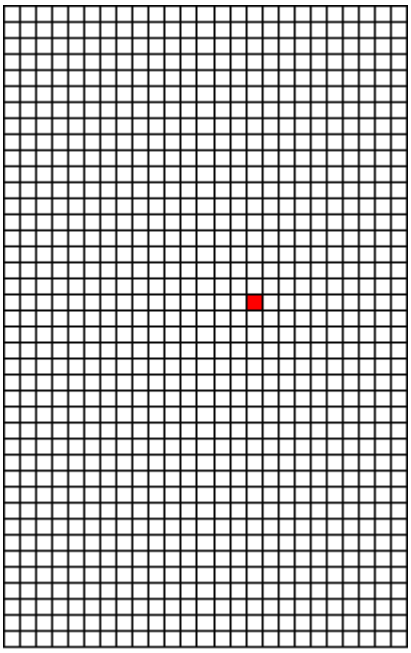
Continue

We want to take a minute to explain how we will describe your risk over these 34 years. Imagine that each small square below represents one person, so that the whole picture represents 1,000 people. RED squares show the people who die over 34 years. WHITE squares show the people who live.

= DEAD

= ALIVE

In the picture below, ONE person out of 1,000 people dies over 34 years.

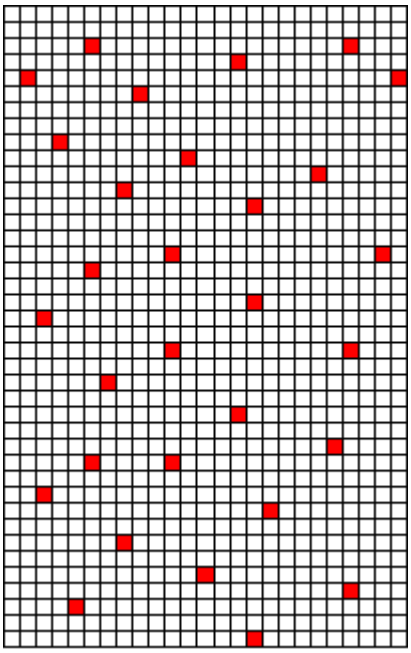


{Form 10 - Private}

Continue

Now imagine that you are one of those 1,000 people in the previous grid. If an illness kills 30 people over the next 34 years, then 970 will NOT have died of this illness at the end of that period. Since you do not know whether this illness will affect you over the next 34 years, we will describe your chance of dying as

30 in 1,000



{Form 11 - Private}

Continue

Next we want to know which illnesses you most want to avoid. We will present you with two illnesses that could affect you. For each illness, we describe how it might affect you.

{Form 12 - Private}

Continue

Consider the possibility that you might experience these two illnesses around these times in your life.

	Respiratory Disease	Colon Cancer
Timeline	Get sick when 65 years old	Get sick when 68 years old

If you have already suffered from one of these illnesses, please view these as possible recurrences.

{Form 13 - Private}

Continue

Each illness may cause pain and disability. Below we describe what it is like to experience moderate and severe pain and disability.

Color key for level of PAIN and DISABILITY:

	Moderate	Pain Some discomfort performing daily activities; most pain can be controlled by medication. Disability Some problems walking, washing, dressing or using the toilet.
	Severe	Pain So bad it impairs daily activities. Difficult to control even with medication. Disability: unable to perform usual daily activities; usually confined to bed; unable to wash, dress, or use toilet independently; unable to communicate well with others.

{Form 14 - Private}

Continue

The pain, disability, and medical treatments associated with these two illnesses would be:

	Respiratory Disease	Colon Cancer
Symptoms / Treatments	No hospitalization Minor surgery Moderate pain for 1 month	1 month of hospitalization Major surgery Severe pain for 18 months Moderate pain for 2 years

{Form 15 - Private}

Continue

If you experience Respiratory Disease or Colon Cancer, it may kill you or you may recover from it. Even if you recover, you may not live until 88 because you are more vulnerable to other illnesses. Assume that these illnesses affect your life expectancy in the following way.

	Respiratory Disease	Colon Cancer
Recovery	Recover at 65	Recover at 71
Life expectancy	Die of something else at 68 instead of 88	Die of something else at 73 instead of 88

Which one shortens your life the most?

Select one answer only

- ◯ Respiratory Disease
- ◯ Colon Cancer
- ◯ Same

{Form 16 - Private}

Next Question

We want to tell you how some new health programs work to reduce your chance of these illnesses.

Like mammograms and prostate exams, these new programs would indicate whether you are at risk for an illness. The big advantage of these new programs is that you and your doctor get better information, much sooner, without uncomfortable procedures.

Your doctor would give you a pin-prick blood test once a year. Each test works by checking for chemicals in your blood that indicate you are at risk for an illness.

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.

Your doctor and the U.S. Food and Drug Administration would certify all of these programs as safe and fully effective before you used them.

{Form 17 - Private}

Continue

We may also ask you about several 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.

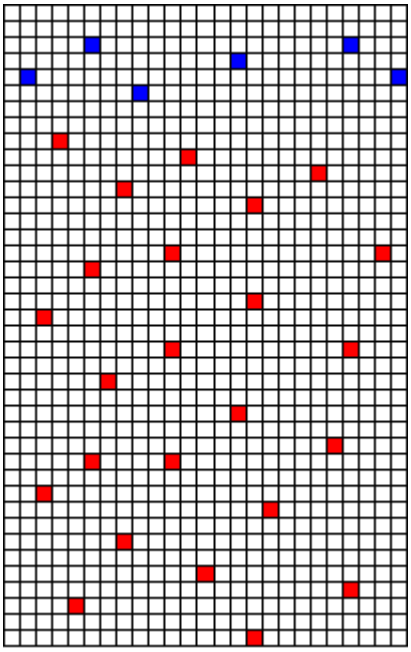
{Form 18 - Private}

Continue

Programs may be very effective at reducing your risk, but you should remember that your risks of dying may be very small.

For example, consider a new program that reduces your risk of dying by 20% - from 30 in 1,000 to 24 in 1,000 - over 34 years. This may sound like a large percentage reduction, but your initial chance of dying was only 30 in 1,000 over the next 34 years. To illustrate this below, the blue squares (■) represent the size of this risk reduction. The red squares (■) represent your remaining chance of dying even with the new program.

20% percent reduction
from 30 in 1,000 to 24 in 1,000



{Form 19 - Private}

Continue

Now we show you how effectively these programs can reduce your chance of respiratory disease and colon cancer. Each program reduces both your risk of getting an illness and your risk of dying from it for the next 34 years.

	Program A for Respiratory Disease	Program B for Colon Cancer
Risk Reduction	75% From 4 in 1,000 to 1 in 1,000	50% From 4 in 1,000 to 2 in 1,000

Which program reduces your risk the most?

Select one answer only

- ◯ Program A for respiratory disease
- ◯ Program B for colon cancer

{Form 20 - Private}

Next Question

Your participation in a program would cost money. These higher costs might take the form of a co-payment when you visit your doctor visit or higher monthly health insurance costs.

To make it easier to compare, we present all costs as monthly costs, and also as annual costs. You would need to pay for, and participate in, a program for the next 34 years to get its benefits.

	Program A for Respiratory Disease	**Program B for Colon Cancer**
Cost to you	$18 per month [= $216 per year]	$4 per month [= $48 per year]

{Form 21 - Private}

Continue

In surveys like this one, people sometimes do not fully consider their future expenses. Please think about what you would have to give up to purchase one of these programs. If you choose a program with too high a price, you may not be able to afford the program when it is offered.

We give you the option to choose "neither program". Some people might choose this option because they:

- cannot afford either program,
- do not believe they face these illnesses or injuries,
- would rather spend the money on other things, or
- believe they will be affected by another illness or injury first.

{Form 22 - Private}

Continue

We explain important points about this table below.

	Program A for Respiratory Disease	Program B for Colon Cancer
Timeline	Get sick when 65 years old	Get sick when 68 years old
Recovery / Life expectancy	Recover at 65 Die of something else at 68 instead of 88	Recover at 71 Die of something else at 73 instead of 88
Risk Reduction	75% From 4 in 1,000 to 1 in 1,000	50% From 4 in 1,000 to 2 in 1,000

We want to be clear about when the benefits from each program begin. For example, the benefits of Program A are that it reduces your risk of respiratory disease from 4 in 1,000 to 1 in 1,000, starting when you are around 65 years old and continuing for the rest of your life. If you DO NOT choose Program A, your risk of respiratory disease will remain at 4 in 1,000 over this time period.

{Form 23 - Private}

Continue

We realize that without proof, you may not accept the idea that these programs are guaranteed to work. Please make your choice as if you have been shown such proof. Remember that all programs would be certified as safe and effective by your doctor and the U.S. Food and Drug Administration.

{Form 24 - Private}

Continue

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 Respiratory Disease	**Program B for Colon Cancer**
Symptoms / Treatment	Get sick when 65 years old No hospitalization Minor surgery Moderate pain for 1 month	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	Recover at 65 Die of something else at 68 instead of 88	Recover at 71 Die of something else at 73 instead of 88
Risk Reduction	75% From 4 in 1,000 to 1 in 1,000	50% From 4 in 1,000 to 2 in 1,000
Costs to you	$18 per month [= $216 per year]	$4 per month [= $48 per year]
Your choice	○ **Reduce my chance of respiratory disease**	○ **Reduce my chance of colon cancer**

○ **Neither Program**

{Form 25 - Private}

Next Question

How difficult was your choice on the previous screen?

Select one answer only

Easy			Somewhat Difficult			Very Difficult
1	2	3	4	5	6	7
○	○	○	○	○	○	○

{Form 26 - Private}

Next Question

You may have chosen Program A, Program B, or neither. Regardless of your choice, we would like to know when, over your lifetime, you think you would first need and benefit from the two programs (if at all).

Your answers below may depend upon the illness or injury in question, as well as your current age, health and family history.

Around when do you think you would begin to value highly the risk reduction benefits of each program?

Select one answer from each column in the grid

	Program A to reduce my chance of respiratory disease	Program B to reduce my chance of colon cancer
For me, benefits would start		
Immediately	○	○
1-5 years from now	○	○
6-10 years from now	○	○
11-20 years from now	○	○
21-30 years from now	○	○
31 or more years from now	○	○
Never (Program would not benefit me)	○	○

{Form 27 - Private}

Next Question

Which reasons best describe why you did not want to pay?

Select all answers that apply

- [ ] I would rather spend the money on something else
- [ ] I did not believe these programs would reduce my risks
- [ ] I will be affected by another illness or injury first
- [ ] I did not believe I faced these health threats
- [ ] I could not afford either program
- [ ] I prefer to take other actions to avoid these risks

{Form 28 - Private}
{Shown only if choice is "neither"}

Next Question

Please evaluate each new pair of programs independently of the ones you saw earlier.

Given the cost, choose the program that reduces the illness you most want to avoid.

Would you prefer Program C, Program D, or neither?

	Program C for Diabetes	**Program D for Stroke**
Symptoms / Treatment	Get sick when 77 years old 6 weeks of hospitalization No surgery Moderate pain for 7 years	Get sick when 65 years old 6 weeks of hospitalization Minor surgery Moderate pain for remaining life
Recovery / Life expectancy	Do not recover Die at 84 instead of 88	Chronic condition Die of something else at 81 instead of 88
Risk Reduction	10% From 10 in 1,000 to 9 in 1,000	20% From 10 in 1,000 to 8 in 1,000
Costs to you	$12 per month [= $144 per year]	$20 per month [= $240 per year]
Your choice	○ **Reduce my chance of diabetes**	○ **Reduce my chance of stroke**

○ **Neither Program**

{Form 29 - Private}

Next Question

How difficult was your choice on the previous screen?

Select one answer only

Easy			Somewhat Difficult			Very Difficult
1	2	3	4	5	6	7
○	○	○	○	○	○	○

{Form 30 - Private}

Next Question

You may have chosen Program C, Program D, or neither. Regardless of your choice, we would like to know when, over your lifetime, you think you would first need and benefit from the two programs (if at all).

Your answers below may depend upon the illness or injury in question, as well as your current age, health and family history.

Around when do you think you would begin to value highly the risk reduction benefits of each program?

Select one answer from each column in the grid

	Program C to reduce my chance of diabetes	Program D to reduce my chance of stroke
For me, benefits would start		
Immediately	○	○
1-5 years from now	○	○
6-10 years from now	○	○
11-20 years from now	○	○
21-30 years from now	○	○
31 or more years from now	○	○
Never (Program would not benefit me)	○	○

{Form 31 - Private}

Next Question

Which reasons best describe why you did not want to pay?

Select all answers that apply

- [ ] I would rather spend the money on something else
- [ ] I did not believe these programs would reduce my risks
- [ ] I will be affected by another illness or injury first
- [ ] I did not believe I faced these health threats
- [ ] I could not afford either program
- [ ] I prefer to take other actions to avoid these risks

{Form 32 - Private}
{Shown only if choice is "neither"}

Next Question

Would you prefer Program E, Program F, or neither?

	Program E for Serious Skin Cancer	Program F for Lung Cancer
Symptoms / Treatment	Get sick when 87 years old 3 days of hospitalization Minor surgery Moderate pain for remaining life	Get sick when 81 years old 6 months of hospitalization Major surgery Moderate pain for 12 months Severe pain for remaining life
Recovery / Life expectancy	Chronic condition Die of something else at 87 instead of 88	Chronic condition Die of something else at 85 instead of 88
Risk Reduction	10% From 30 in 1,000 to 27 in 1,000	20% From 30 in 1,000 to 24 in 1,000
Costs to you	$19 per month [= $228 per year]	$50 per month [= $600 per year]
Your choice	○ **Reduce my chance of serious skin cancer**	○ **Reduce my chance of lung cancer**

○ **Neither Program**

{Form 33 - Private}

Next Question

How difficult was your choice on the previous screen?

Select one answer only

Easy			Somewhat Difficult			Very Difficult
1	2	3	4	5	6	7
○	○	○	○	○	○	○

{Form 34 - Private}

Next Question

Around when do you think you would begin to value highly the risk reduction benefits of each program?

Select one answer from each column in the grid

	Program E to reduce my chance of serious skin cancer	Program F to reduce my chance of lung cancer
For me, benefits would start		
Immediately	○	○
1-5 years from now	○	○
6-10 years from now	○	○
11-20 years from now	○	○
21-30 years from now	○	○
31 or more years from now	○	○
Never (Program would not benefit me)	○	○

{Form 35 - Private}

Next Question

Which reasons best describe why you did not want to pay?

Select all answers that apply

- [ ] I would rather spend the money on something else
- [ ] I did not believe these programs would reduce my risks
- [ ] I will be affected by another illness or injury first
- [ ] I did not believe I faced these health threats
- [ ] I could not afford either program
- [ ] I prefer to take other actions to avoid these risks

{Form 36 - Private}
{Shown only if choice is "neither"}

Next Question

Would you prefer Program G, Program H, or neither?

	Program G for Alzheimer's Disease	Program H for Heart Disease
Symptoms / Treatment	Get sick when 65 years old 4 years of long-term care No surgery Moderate disability for 4 years	Get sick when 71 years old 2 weeks of hospitalization No surgery Moderate pain for remaining life
Recovery / Life expectancy	Do not recover Die at 69 instead of 88	Chronic condition Die of something else at 86 instead of 88
Risk Reduction	10% From 40 in 1,000 to 36 in 1,000	5% From 40 in 1,000 to 38 in 1,000
Costs to you	$19 per month [= $228 per year]	$15 per month [= $180 per year]
Your choice	○ **Reduce my chance of Alzheimer's disease**	○ **Reduce my chance of heart disease**

○ **Neither Program**

{Form 37 - Private}

Next Question

How difficult was your choice on the previous screen?

Select one answer only

Easy			Somewhat Difficult			Very Difficult
1	2	3	4	5	6	7
○	○	○	○	○	○	○

{Form 38 - Private}

Next Question

Around when do you think you would begin to value highly the risk reduction benefits of each program?

Select one answer from each column in the grid

	Program G to reduce my chance of Alzheimer's disease	Program H to reduce my chance of heart disease
For me, benefits would start		
Immediately	○	○
1-5 years from now	○	○
6-10 years from now	○	○
11-20 years from now	○	○
21-30 years from now	○	○
31 or more years from now	○	○
Never (Program would not benefit me)	○	○

{Form 39 - Private}

Next Question

Which reasons best describe why you did not want to pay?

Select all answers that apply

- [ ] I would rather spend the money on something else
- [ ] I did not believe these programs would reduce my risks
- [ ] I will be affected by another illness or injury first
- [ ] I did not believe I faced these health threats
- [ ] I could not afford either program
- [ ] I prefer to take other actions to avoid these risks

{Form 40 - Private}
{Shown only if choice is "neither"}

Next Question

This is the final pair of programs.

Would you prefer Program I, Program J, or neither?

	Program I for Traffic Accident	Program J for Heart Attack
Symptoms / Treatment	Suffer injury when 73 years old No hospitalization No surgery Severe pain for a few hours	Get sick when 67 years old No hospitalization No surgery Severe pain for a few hours
Recovery / Life expectancy	Do not recover Die suddenly at 73 instead of 88	Do not recover Die suddenly at 67 instead of 88
Risk Reduction	5% From 40 in 1,000 to 38 in 1,000	10% From 40 in 1,000 to 36 in 1,000
Costs to you	$4 per month [= $48 per year]	$17 per month [= $204 per year]
Your choice	○ **Reduce my chance of traffic accident**	○ **Reduce my chance of heart attack**

○ **Neither Program**

{Form 41 - Private}

Next Question

How difficult was your choice on the previous screen?

Select one answer only

Easy			Somewhat Difficult			Very Difficult
1	2	3	4	5	6	7
○	○	○	○	○	○	○

{Form 42 - Private}

Next Question

Around when do you think you would begin to value highly the risk reduction benefits of each program?

Select one answer from each column in the grid

	Program I to reduce my chance of traffic accident	Program J to reduce my chance of heart attack
For me, benefits would start		
Immediately	○	○
1-5 years from now	○	○
6-10 years from now	○	○
11-20 years from now	○	○
21-30 years from now	○	○
31 or more years from now	○	○
Never (Program would not benefit me)	○	○

{Form 43 - Private}

Next Question

Which reasons best describe why you did not want to pay?

Select all answers that apply

- [ ] I would rather spend the money on something else
- [ ] I did not believe these programs would reduce my risks
- [ ] I will be affected by another illness or injury first
- [ ] I did not believe I faced these health threats
- [ ] I could not afford either program
- [ ] I prefer to take other actions to avoid these risks

{Form 44 - Private}
{Shown only if choice is "neither"}

Next Question

Do you tend to put more effort into protecting your health now than you did ten years ago?

Select one answer only

- Much more
- Somewhat more
- About the same
- Somewhat less
- Much less

{Form 45 - Private}

Next 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?

Select one answer only

- Very likely
- Somewhat likely
- Somewhat unlikely
- Very unlikely

{Form 46 - Private}

Next Question

Did you consider whether you could actually afford to pay for these programs over your lifetime?

Select one answer only

- Yes
- Somewhat
- No

{Form 47 - Private}

Next 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?

Select one answer only

- highly confident
- somewhat confident
- not at all confident

{Form 48 - Private}

Next Question

We cannot perfectly predict how long we will live. But based on our health and family history, most of us have some idea about how long we might live.

Until what age do you expect to live? Please check your best guess.

Select one answer only

54	65	76	87	97
55	66	77	88	98
56	67	78	89	99
57	68	79	90	100
58	69	80	91	101
59	70	81	92	102
60	71	82	93	103
61	72	83	94	104
62	73	84	95	105
63	74	85	96	More than 105
64	75	86		

{Form 49 - Private}

Next Question

Thank you for your time!

{Form 50 - Private}

Finish