The Journal of Risk and Uncertainty, 24:2; 161–186, 2002

Age, Health and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents

ALAN KRUPNICK
Resources for the Future

ANNA ALBERINI
University of Maryland

MAUREEN CROPPER*
University of Maryland and World Bank

NATHALIE SIMON
US Environmental Protection Agency

BERNIE O'BRIEN
McMaster University

RON GOEREE
McMaster University

MARTIN HEINTZELMAN
Resources for the Future

Abstract

We present the results of a contingent valuation survey eliciting willingness to pay (WTP) for mortality risk reductions. The survey was self-administered using a computer by 930 persons in Hamilton Ontario aged 40 to 75. Visual and audio aides were used to enhance risk comprehension. Mean WTP figures for a contemporaneous risk reduction imply a value of a statistical life of approximately C$l.2 to C$3.8 million (1999 C$). Mean WTP is constant with age up to 70 years, and is about 30 percent lower for persons aged 70 and older. WTP is unaffected by physical health status, but is affected by mental health.

Keywords: value of a statistical life, mortality risks, benefit-cost analysis

JEL Classification: I18, Q28

* Address for correspondence: Maureen Cropper, The World Bank, MSN MC 2-205, 1818 H Street NW, Washington, DC 20433; e-mail: mcropper@worldbank.org

Much of the justification for environmental rulemaking rests on estimates of the benefits to society of reduced mortality rates. Reductions in risk of death are arguably the most important benefit underlying many of the United States Environmental Protection Agency's (USEPA) legislative mandates, including the Safe Drinking Water Act, the Resource Conservation and Recovery Act and the Clean Air Act. For example, in two recent analyses of the benefits of U.S. air quality legislation, *The Benefits and Cost of the Clean Air Act, 1970–1990* (USEPA, 1997) and *The Benefits and Cost of the Clean Air Act, 1990–2010* (USEPA, 1999), over 80 percent of monetized benefits were attributed to reductions in premature mortality. These benefits are equally important in environmental cost-benefit analyses performed in Canada (Canada-Wide Standards Development Committee for PM and Ozone, 1999).

There are two sources of empirical estimates of individuals' willingness to pay (WTP) for mortality risk reductions: revealed preference studies, based on compensating wage data or consumer behavior, and stated preference studies, including those employing contingent valuation methods. From the perspective of valuing lives saved by environmental programs both estimation techniques—as applied to date—share a common shortcoming. They focus on measuring the value that prime-aged adults place on reducing their risk of dying, whereas the majority of statistical lives saved by environmental programs, according to epidemiological studies, appear to be the lives of older people and people with chronically impaired health. It has been conjectured that older people should be willing to pay less for a reduction in their risk of dying than younger people on the grounds that they have fewer expected life years remaining. Theory, however, cannot predict exactly how WTP varies with age, and, to our knowledge, few empirical studies have been conducted that include subjects over the age of 65. Likewise, there are no studies that examine the impact of health status on WTP for mortality risk changes.

The goal of this research is to estimate what older people are willing to pay to reduce their risk of dying, and to examine the impact of current health status on WTP. We accomplish this through a contingent valuation survey that is administered to persons 40 to 75 years old. Targeting this age range allows us to examine the impact of age on WTP, thus providing an empirical answer to the above speculations, and allows us to compare our WTP estimates with those from previous studies. We measure health status in two ways. Respondents are asked whether they have ever been diagnosed as having one of several chronic heart or lung diseases, or cancer. To further capture the severity of the disease (or other chronic health conditions) we ask respondents to complete a detailed health questionnaire, Short Form 36 (Ware, Kosinski and Keller, 1997), which has been shown to correlate well with severity of various chronic illnesses (Bousquet et al., 1994).

This paper reports the results of our survey, which was self-administered using a computer by 930 persons in Hamilton, Ontario in 1999. The survey uses audio and visual aids to communicate both baseline risk of death and risk changes. Respondents are given experience with graphical representations of risks of death (depicted by colored squares on a rectangular grid) and are tested for comprehension of probabilities before being asked WTP questions. We credit these efforts at risk communication with the fact that mean WTP of respondents faced with larger risk reductions exceeds mean WTP of respondents faced with smaller risk reductions—i.e., our respondents pass external scope tests.

Our mean WTP estimates for a contemporaneous risk reduction imply a value of a statistical life (VSL) of approximately C$1.2 to C$3.8 million (1999 C$), or from U.S.$0.96 million to U.S.$3.04 million (1999 U.S.$),[1] depending on the size of the risk change valued. These estimates are in the range of estimates from revealed preference studies (Moore and Viscusi, 1988; Kniesner and Leeth, 1991; Cousineau, Lacroix and Girard, 1992) and from previous contingent valuation studies (Gerking, DeHaan and Schulze, 1988; Jones-Lee, Hammerton and Philips, 1985; Miller and Guria, 1991; Viscusi, Magat and Huber, 1991).

Interestingly, we find that age has no effect on WTP until roughly age 70: Mean WTP is similar across age groups up to about 70 years of age, and is about 30 percent lower for persons aged 70 and older. Regarding physical health, dummy variables for individual chronic diseases are not significantly related to WTP, with the exception of cancer. We find that the WTP of the representative person with cancer is substantially (64%) larger (at an 8% confidence level) than a person with the same personal characteristics but without cancer. We caution, however, that this result is based on only 26 respondents with cancer, who were well enough to come to the facility where the survey was administered.

Physical health status, as measured by SF-36, has no effect on WTP. The index of physical functioning computed from SF-36, which measures limitations in everyday activities caused by illness, has no statistically significant impact on WTP. Other components of physical health captured by SF-36 (e.g., the energy/vitality score, general health perception score) also have no impact on WTP. By contrast, the SF-36 mental health score, which measures symptoms of psychological distress, *is* significantly related to WTP. People with fewer symptoms of psychological distress are willing to pay significantly more to reduce their chance of dying.

The remainder of the paper is organized as follows. In Section 1, we discuss the nature and limitations of current estimates of VSL. In Section 2, we describe our goals and how we structured the survey questionnaire to achieve them. Section 3 presents the questionnaire, Section 4 our protocols, Section 5 the results, and Section 6 concluding remarks.

1. The value of reductions in mortality risks: The state of the art

1.1. The nature of mortality risk reductions from environmental programs

Life saving benefits from environmental regulations have been quantified for the conventional air pollutants, especially particulate matter, and for carcinogens. These studies suggest that life-saving benefits are concentrated among persons 65 years of age and older and may disproportionately benefit people with pre-existing chronic conditions. Other health and safety regulations, such as those intended to reduce foodborne pathogens, also disproportionately benefit older persons and persons in compromised health. Epidemiological evidence for the link between older people and air pollution comes from two directions. First, epidemiological studies typically assume that the effect of a change in pollution concentrations is proportional to baseline mortality rates.[2] Since death rates are higher for older persons, this implies that the benefits of reducing exposure to air pollution accrue primarily to older people. Based on Pope et al. (1995), the EPA (1997) estimates that three-quarters of the statistical lives saved by the Clean Air Act in 1990 as a result of reducing particulate matter are persons 65 years of age and older. Second, epidemiological studies have found larger changes in mortality rates for people over 64 than for younger people (Schwartz 1991, 1993).

Reducing exposure to pollution may also reduce risk of cancer.[3] Although the toxicological studies that are used to quantify cancer risks provide only an estimate of *lifetime* cancer risk, rather than age-specific risk estimates, it is reasonable to assume that the age distribution of deaths from environmentally induced cancers follows the same pattern as cancer mortality rates from all causes. Since cancer mortality rates are concentrated among individuals aged 65 and over, the statistical lives saved by reducing exposure to carcinogens will be concentrated among people in the same age group.[4]

Epidemiological studies also suggest that persons with chronic heart or lung conditions are likely to benefit disproportionately from improvements in air quality. For example, Schwartz (1991), Schwartz and Dockery (1992), and Pope et al. (1995) find that changes in particulate concentrations have a larger impact on deaths due to cardiovascular disease and chronic obstructive lung disease than on all deaths. This has caused some observers to suggest that the value of lives saved by air pollution should reflect the compromised health of the beneficiaries (EOP Group, Inc., 1997). It is not, however, clear that people with chronic heart and lung disease would pay less than healthier individuals to reduce their risk of dying.

1.2. Current approaches to valuing mortality risk reductions

In benefit-cost analyses of health and safety regulations, including environmental regulations, it is standard practice to ignore the health status of people whose lives are extended by the regulation. The age of persons saved is sometimes incorporated by converting the value of a statistical life from a labor market study (or other source) into a value per life-year saved.[5] The value of a life-year can then be multiplied by discounted remaining life expectancy to value the statistical lives of persons of different ages. This procedure is, however, *ad hoc*. It assumes that the value per life-year saved is independent of age, and it is sensitive to the rate used to discount the value of future life-years, which is usually assumed by the researcher rather than estimated on the basis of actual behavior.[6]

Evidence from contingent valuation studies (Jones-Lee, Hammerton and Philips, 1985) suggests that willingness to pay is not proportional to remaining life expectancy; however, policymakers may be reluctant to rely on such studies unless it can be demonstrated that they pass tests of internal and external validity. One measure of the success of a contingent valuation survey is that, when different groups of respondents are asked to value risk changes of different magnitudes, WTP increases with the size of the risk change. An external scope test is passed when the mean WTP of respondents faced with the larger risk change is significantly greater than the mean WTP of the respondents faced with the smaller risk change. An internal scope test is passed when a *respondent's* WTP increases with the size of the risk reduction. In the context of valuing risk changes, however, a more stringent criterion can be applied. If respondents maximize expected utility or, more generally, if their utility function is linear in probabilities, WTP for small risk changes should increase *in proportion to* the size of the risk change.

As a recent literature review by Hammitt and Graham (1999) demonstrates, few contingent valuation studies of mortality risks pass either internal or external scope tests. In some cases (e.g., Jones-Lee, Hammerton and Philips, 1985; Smith and Desvousges, 1987) WTP fails to increase significantly with the size of the risk change. Only three contingent valuation studies designed to value mortality risks pass external scope tests. All of these studies were conducted in the context of traffic safety and two involved extremely small samples ($N < 110$). None of these studies focused on valuing mortality risk reductions among older people and none examined the impact of health status on WTP for risk reductions.

2. Valuing mortality risks among older persons

2.1. Goals of the survey

The goal of our survey is to estimate what older people would pay for a reduction in their risk of dying and to examine the impact of health status on willingness to pay. We target a population ranging in age from 40 (the mean age of workers in compensating wage studies) to 75 and collect extensive information on health status.[7]

We ask respondents to value annual risk reductions on the order of 10^{-4}. Risk changes valued in labor market studies are on the order of 1 in 10,000 per year. A risk change of this order of magnitude could also be delivered by an environmental program (e.g., air pollution control). For instance, the Pope et al. study (1995) predicts that a 10 $\mu g/m^3$ change in PM10 results in an annual average change in risk of death of 2.4 in 10,000, while studies based on time series generally predict that the same change in pollution levels results in a 0.8 in 10,000 risk change.

For use in benefit-cost analyses, it is important that risk reductions be a private good; i.e., that we estimate each respondent's WTP to reduce his *own* risk of dying. For this reason, we have chosen an abstract product (not covered by health insurance) as the mechanism by which risk reductions are delivered. In practice, most environmental programs reduce mortality risks for all persons in an exposed population: in other words, risk reductions are a public good. Johansson (1994) and Jones-Lee (1991) have shown, however, that when people exhibit pure altruism, maximization of net social benefits calls for equating the sum of individuals' marginal willingness to pay to reduce risks to themselves to the marginal cost of the risk reductions.[8] Therefore, the appropriate measure of benefits is the sum of private WTP for reductions in risk.

2.2. Avoiding past pitfalls

The failure of many CV studies to pass tests of internal and external validity may be traced to three types of problems.

1. Respondents may not understand the risk changes they are asked to value. Our survey relies on a graph containing 1,000 squares to communicate probability of dying. White squares denote chances of surviving, red squares represent chances of dying. Reductions in the risk of dying are represented by changing red squares to blue.

Because we value *annual* risk changes on the order of 10^{-4}, the graph represents the chances of dying (surviving) over a 10-year period with risks on the order of 10^{-3}. The use of a 10-year period makes it possible to represent risks using 1,000 squares. In our questionnaire development, we found that respondents regarded grids with more squares (e.g., 10,000 or 100,000) as confusing, and tended to dismiss such small risk changes as insignificant.

Each respondent goes through the survey on a computer screen, at his own pace. To improve understanding of risk changes, we encourage respondents to think about changes in mortality risks by showing them side-by-side depictions of the risks with and without the product, and by asking them questions to test their understanding of how risks (and risk changes) are represented. If the respondent answers a question incorrectly, he is provided additional educational information and is asked an additional, similar question.

2. Respondents may not believe that the risk changes (or baseline risks) apply to them. We found that when respondents are told their baseline risk of dying over the next 12 months, they often believe that the risks do not apply to them. In focus groups, respondents more readily accepted baseline risks over longer periods, which was another reason for using the 10 year risk reduction period. We also asked respondents to assume the risks were their own.

3. Respondents may lack experience in trading money for quantitative risk changes or fail to realize that they engage in this activity. We acquaint respondents with the *quantitative* risk reductions associated with medical tests and products with which the respondent may be familiar (e.g., mammograms, colon cancer screening tests, medicine to reduce blood pressure) prior to asking what he or she would pay for a product that will reduce risk of dying. In doing so, we keep the cost information provided to the respondent qualitative in nature.

Regarding the timing of payments, focus groups and one-on-one interviews suggested that payments for risk reductions should be made annually, over a ten-year period. We use graphs to convey the timing of the payments and the relationship between the timing of payments and risk reductions.

3. Description of the questionnaire

The survey instrument used in this project was developed over a period of several years. The development work included extensive one-on-one interviews in the U.S., pre-tests in the U.S. and Japan, and several focus groups in Hamilton, Ontario, including one at a senior citizen recreation center, followed by another pre-test.

The survey was administered on a computer with a simplified keypad, which was color-coded and especially labeled for use with the survey (e.g., “Press the BLUE key to see the next screen”). Respondents moved through the survey at their own pace. Words on each screen appeared in large font, and were read to the respondent by a voice-over. We chose a self-administered computer questionnaire to avoid interviewer effects, deliver high-quality graphics, and accommodate age- and gender-specific risks and follow-up questions (see Figure 1).

The questionnaire is divided into five parts. Part I elicits personal information, including health information about the respondent and his immediate family. Part II introduces the subject to simple probability concepts using coin tosses and roulette wheels, working up to the 1,000-square grid. He is asked to distinguish which of two grids shows the higher risk of death and which person he would rather be. The baseline risk of death for a person of the respondent’s age and gender is then presented both numerically and graphically.

Part III presents each respondent with the leading causes of death for someone of his age and gender, examples of common risk-mitigating behaviors together with the quantitative risk reductions they achieve, and a qualitative estimate of the costs associated with them (“inexpensive,” “moderate,” and “expensive”).[9] Part IV elicits WTP for risk reductions of a given magnitude, occurring at a specified time, using dichotomous choice methods. Respondents are randomly assigned to one of two sub-samples. As shown in Table 1, respondents in one sub-sample (Wave 1) are first asked if they

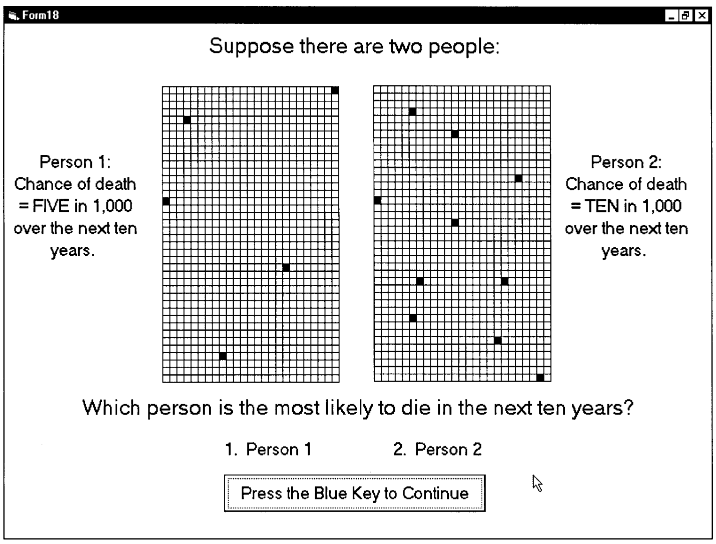

Figure 1. Use of grids to represent probabilities in mortality risk questionnaire.

Table 1. Survey design

Group of respondents	Initial risk reduction valued	Second risk reduction valued	Future risk reduction valued
Wave 1 ($N = 630$)	5 in 1,000	1 in 1,000	5 in 1,000
Wave 2 ($N = 300$)	1 in 1,000	5 in 1,000	5 in 1,000

are willing to pay for a product that, when used and paid for over the next ten years, will reduce baseline risk by 5 in 1,000 over the 10-year period (WTP5), i.e., by 5 in 10,000 annually. In the second WTP question, risks are reduced by 1 in 1,000 (WTP1), i.e., by 1 in 10,000 annually. Respondents in the second sub-sample (Wave 2), are given the 1 in 1,000 risk change question first. The risk reductions delivered by the products are shown by changing the appropriate number of baseline squares from red to blue.[10]

Table 2 provides the bid structure used in the survey. All respondents are asked follow-up dichotomous choice questions to obtain more information about WTP. All WTP dichotomous choice questions answered by "No-No" responses are followed by a question asking if the respondent is willing to pay anything at all, and if so, how much. Respondents are then asked, on a 1 to 7 scale, their degree of certainty about their responses.

Part V asks an extensive series of debriefing questions, followed by some final sociodemographic questions (e.g., education and household income). The debriefing questions are used to identify respondents who had trouble comprehending the survey or did not accept the risk reduction being valued.

The computerized questionnaire is followed by a pencil-and-paper 36-question Quality of Life survey (Short Form-36, abbreviated SF-36), which is used routinely in the medical community to gauge physical functionality, and mental and emotional health states (Ware, Kosinski and Keller, 1997). The 36 health questions supplement those posed at the beginning of the interview and may be used to construct eight indices commonly used in the health literature.

Table 2. Bid structure in the mortality risk survey (1999 C$)

Group of respondents	Initial payment question	Follow-up question (if "yes")	Follow-up question (if "no")
I	100	225	50
II	225	750	100
III	750	1100	225
IV	1100	1500	750

4. Survey protocols

Our survey was administered to 930 subjects in Hamilton, Ontario in 1999, over a five-month period. Subjects were recruited by telephone through random digit dialing and asked to take the survey at a facility in downtown Hamilton. Because of the need to travel to a centralized facility, response rates were low. Out of 17,841 residential phone contacts 8,260 were "cooperative," but 4,917 households proved ineligible for age reasons (there was no one in the household between ages 40 to 75). Among the 3,591 eligible households, 455 declined to participate because of mobility problems and 1,079 refused, stating that the incentive payment (C$35) was insufficient. 1,545 persons agreed to participate in the survey, but in fact only 930 (60%) kept their appointments. All persons who began the survey completed it, bringing our response rate to 26 percent.[11]

Because our target population was persons between 40 and 75 years of age, one-third over age 60, and because we asked people to go to a centralized facility, we cannot make claims that our sample is representative of the population of Hamilton. Nevertheless, our sample is very similar to the population of Hamilton aged 40 to 75 in many respects, in particular with respect to the prevalence of chronic conditions. According to the National Population Health Survey, 8.9 percent of the Hamilton population suffered from heart disease in 1998 and 15.2 percent had high blood pressure. Similarly, 2.6 percent of the population had bronchitis or emphysema while 8.3 percent suffered from asthma. A total of 2.0 percent of the population aged 40 to 75 years reported having cancer (Stieb, personal communication).

Table 3 provides descriptive statistics for our entire sample and for each subsample. The average age of respondents was 54 years old, with 31 percent of the sample above age 60 and 9 percent above age 70, which is consistent with our sampling plan. Although 80 percent of the sample completed high school, only 20 percent completed a university degree. (Average educational attainment was 13.7 years of schooling.) The average household income in our sample was approximately C$58,000 (1999 C$), with median household income equal to C$50,000 (1999 C$). The average individual in our sample is thus slightly wealthier and more educated than the average Hamilton resident, but is very similar to the average Ontario resident.

Although Canada has universal health insurance, many people carry supplemental insurance. In our sample, 32 percent of respondents reported having additional insurance. Most respondents rated their health as very good to excellent, although 41 percent reported having some form of chronic respiratory or heart disease. Only 4 percent had been hospitalized in the last five years.

Table 3 also reports the average index scores from answers to the SF-36 questionnaire for physical functioning and mental health.[12] While Canadian index scores for SF-36 are currently not available for comparison, the results suggest that for physical functioning our respondents (who score on average 84 out of 100) are slightly below the U.S. national average, and for mental health (77 out of 100) they are slightly above it.

Table 3. Characteristics of respondents

	Mean (std. dev.)		
Variable name	Wave 1 ($N = 630$)	Wave 2 ($N = 300$)	Total sample ($N = 930$)
Age in years	54.0 (10.2)	54.4 (10.4)	54.2 (10.3)
Female	53 percent	56 percent	54 percent
Years of education	13.7 (3.02)	13.6 (3.17)	13.7 (3.07)
Annual household income (C$)	$58,896 ($35,596)	$57,446 ($33,103)	$58,425 ($35,487)
Has supplemental insurance	32.0 percent	31.4 percent	32.0 percent
Mental health score (1–100) (Bipolar scale measuring mental health status; Midrange score indicates respondent reports no symptoms of psychological distress; Score = 100 indicates respondent reports frequently feeling happy, calm and peaceful.)	76.5 (16.8)	77.5 (15.6)	76.8 (16.4)
Physical functioning score (1–100) (Measures limitations in behavioral performance of everyday physical activities. Low score indicates limitations in performing all physical activities due to health; High score indicates that all types of physical activities are performed with no limitations due to health)	81.1 (21.2)	82.2 (21.8)	81.4 (21.4)
Baseline risk over 10 years (times 1,000)	122.6 (123.5)	122.5 (119.4)	122.6 (122.1)
Heart disease	10.3 percent	8.3 percent	9.6 percent
High blood pressure	19.4 percent	21.8 percent	20.2 percent
Cancer	4.1 percent	2.0 percent	3.4 percent
Asthma	11.3 percent	8.4 percent	10.3 percent
Bronchitis, emphysema, or chronic cough	14.1 percent	13.7 percent	14.0 percent
Self-assessed years before death	26.0 (12.5)	26.2 (12.5)	26.1 (12.4)

5. Results

5.1. Risk comprehension

The statistics in Table 4 refer to questions testing comprehension of risks, listed in order from the weakest to the most stringent criteria. The probability *test* question asks respondents which person, 1 or 2, has the higher risk of death, when risks are represented using darkened squares on a grid. The probability *choice* question asks respondents which one of these people *they would rather be*. While 11.5 percent and 13 percent of the respondents chose the "wrong person" in these questions, only 2.6 percent of the sample failed both tests. About one percent of the sample confirmed that they preferred to be the person with the higher risk when prompted to confirm or change their answer to the probability choice question. In a question at the end of the survey, 7 percent of respondents assessed their understanding of chance as fair to poor.

Table 4. Descriptive statistics: Risk comprehension

Percent of respondents who...	Wave 1 ($N = 630$)	Wave 2 ($N = 300$)	Total sample ($N = 930$)
Chose wrong person in probability test	11	13	12
Chose wrong person in the probability choice	12	16	13
Chose wrong person in probability test and wrong person in the probability choice	2	3	3
Confirmed wrong person in probability choice	1	2	1
Selected 3 or less in self-assessed understanding of probability (1 = does not understand at all; 5 = has excellent understanding of probability)	7	7	7

5.2. Understanding/acceptance of the scenario

Table 5 reports statistics on the percent of respondents who questioned assumptions made in the survey or said they did not understand the scenario presented to them. On debriefing, 20 percent of respondents did not think that the baseline risk of death given in the survey applied to them. Thirty-one percent of the sample did not believe our abstract product would work, but only 20 percent said this affected their WTP, with most (17 percent) saying they bid less because of it. Side effects were expected by 25 percent of the sample, an effect that may potentially lower WTP. On the other hand, 40 percent of the sample thought the product would provide benefits in addition to reducing mortality risk, and 20 percent of the sample stated that these other benefits increased their WTP.

Regarding the payment for the life-extending product, 13 percent of the sample professed not to understand the payment scheme, while 26 percent of the sample said they did not think about affordability when they answered. As will become clear below, this response does not mean that subjects ignored their income constraints and reported excessively large WTP amounts. Instead, we found that this type of response is very common among people who are not willing to pay anything at all for a product, implying that

Table 5. Descriptive statistics: Scenario acceptance

Percent of respondents who...	Wave 1 ($N = 630$)	Wave 2 ($N = 300$)	Total sample ($N = 930$)
Did not believe the stated risks applied to them (Flag7)	19	22	20
Had doubts about the product's effectiveness	30	31	31
Had doubts about the product's effectiveness, and said doubts affected WTP	20	20	20
Thought product might have side-effects	26	23	25
Thought about other benefits of the product (Flag11)	40	38	40
Said other benefits influenced WTP	21	19	20
Did not understand the payment scheme (Flag16)	12	15	13
Did not consider whether they could afford the payment (Flag15)	23	27	26

most people who so answered this debriefing question did not bother to consider income constraints because they had no intention of buying the product.

We believe that it is an empirical matter whether beliefs about product efficacy, side effects or non-mortality benefits, and perceptions about the risk reduction scenario affect WTP. We therefore investigate the impact of these responses below by including them as regressors in an equation to explain WTP.

5.3. WTP responses: Current risk reductions

When respondents are asked to value several "commodities" within an interview, WTP tends to be affected by the order in which the commodities are presented. Our survey is not exempt from this effect; hence we analyze WTP for only the first risk reductions presented in the questionnaire.[13] Therefore, for respondents in wave 1 and wave 2 we examine WTP for risk reductions equal to 5 in 1,000 and 1 in 1,000, respectively.

We depict the proportion of "yes" responses to the initial payment questions in Figure 2. The figure indicates that in both waves the proportion of "yes" responses falls for higher initial bids, and that the proportion of "yes" responses is higher at all bids for the 5/1,000 risk change than for the 1/1,000 risk change. This suggests that responses are consistent with economic theory, and that WTP is sensitive to the magnitude of the risk reduction.[14, 15]

Respondents who said "no" to both initial and follow-up bids were asked whether they were willing to pay anything at all for the product delivering the risk reduction. The fraction of respondents refusing to pay for the product was 19.5 percent for the 5 in 1,000 risk reduction (wave 1) and 36.8 percent for the 1 in 1,000 risk reduction from

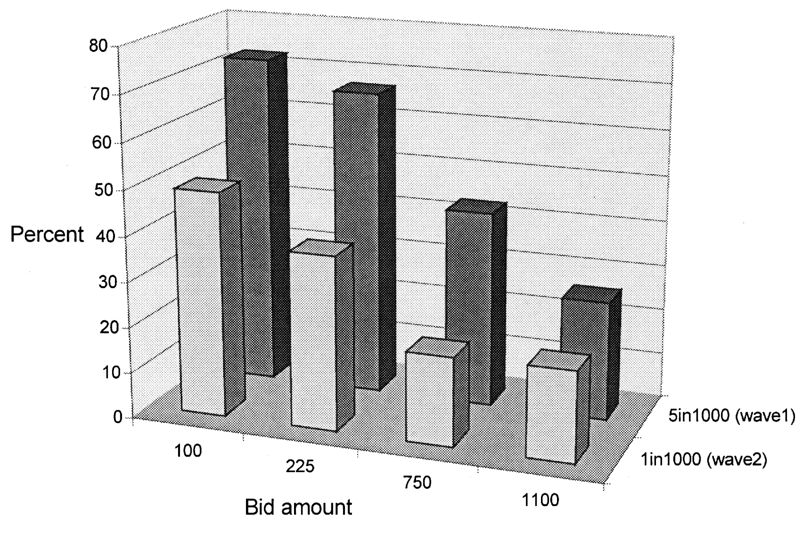

Figure 2. Percentage of "yes" responses to the initial payment question.

wave 2. While it is comforting that the percent of zeros is smaller for the larger risk reduction, these numbers are large, but not very different from those seen in previous CV studies about risk reduction.

To identify reasons for not paying anything at all, we estimated a probit model to explain zero WTP for the 5 in 1,000 risk reduction for wave 1 (shown in the Appendix, Table A.2). As Table A.2 indicates, there is no particular association between zero WTP and individual characteristics, with the exception of the mental health score from SF-36. People with lower mental health scores are more likely to hold zero WTP values for the risk reduction. People with zero WTP decline to pay anything at all for the product without even considering whether they can afford the payments.

Because zero WTP values are numerous, the appropriate statistical framework to model WTP is an interval data/continuous data variant of the tobit model, also known as the "spike" model (Kriström, 1997). We adapt the tobit/"spike" model to the present sample, which includes a mix of zero WTP, continuous observations, and interval data, assuming a normal distribution for WTP. The contributions to the likelihood are:

$$\frac{1}{\sigma}\phi\left(\frac{\mathrm{WTP}_i - \mu}{\sigma}\right) \quad \text{for positive WTP observed on a continuous scale;} \qquad (1)$$

$$\Phi\left(\frac{0 - \mu}{\sigma}\right) \quad \text{for those respondents who are not willing to pay} \qquad (2)$$

anything at all for the product; and

$$\Phi\left(\frac{\mathrm{WTP}_i^H - \mu}{\sigma}\right) - \Phi\left(\frac{\mathrm{WTP}_i^L - \mu}{\sigma}\right) \quad \text{for all other respondents,} \qquad (3)$$

where ϕ and Φ denote the standard normal pdf and cdf, respectively. Mean WTP is computed by integrating the survival function (i.e., one minus the cdf), which displays a "spike" at zero, with respect to the bid from zero to infinity (Kriström, 1997).

Mean WTP for a 5 in 1,000 risk reduction taking place over the next 10 years is C$601 (1999 C$) per year (standard error C$28). The corresponding figure for the 1-in-1,000 risk reduction is C$368 per year (s.e. C$29). A Wald test shows that mean WTP for the 5 in 1,000 risk reduction *is* statistically greater than mean WTP for the 1 in 1,000 risk reduction (Wald statistic = 33.64; p-value < 0.0001). In other words, our estimates pass the external scope test, and prove sensitive to the size of the risk reduction. However, WTP is *not* proportional to the size of the risk reduction: Mean WTP for the 5 in 1,000 risk reduction is much less than 5 times mean WTP for the 1 in 1,000 risk reduction (Wald statistic = 72.27, p-value < 0.0001).[16]

The WTP figures can be used to compute the corresponding value of a statistical life (VSL). We computed VSL by dividing annual WTP by the size of the annual risk reduction (5 in 10,000 or 1 in 10,000).[17] Since the hypothesis of proportionality is violated, VSL changes with the size of the risk reduction, and is greater when WTP for the 1/10,000 risk reduction is used in the calculation. The respective VSLs are C$1,202,000 and C$3,684,000. These figures are below or equal to the age-adjusted VSL used by Health Canada and well below the official C$7.5 million figure used by USEPA.[18]

5.4. *Understanding of risks and* WTP

It should be noted that our WTP figures include in the sample used for estimation *all* respondents, regardless of their performance on our probability tests, provided that they do not report open-ended WTP amounts that exceed the lowest bid amount they were offered. This basic data cleaning reduced the sample size from 630 to 625 (for wave 1) and from 300 to 298 (for wave 2).

In Table 6, we contrast these WTP figures with those obtained after respondents who failed some of the probability comprehension tests are excluded from the usable sample. The second row of Table 6 shows the estimates of mean WTP obtained after further dropping respondents who failed to correctly identify the person with the higher risk of death from the graphs presented at the beginning of the survey *and* who insisted that they would rather be the person with the higher risk of death (i.e., persons with Flag1 = 1). This further reduces the size of wave 1 to 611 respondents and wave 2 to 288 respondents, and is our preferred data cleaning criterion. Use of the cleaned sample

Table 6. Sensitivity of mean WTP for current risk reductions to the data cleaning criteria—spike model

Data cleaning criteria	Mean WTP for the 5 in 1,000 risk reduction (s.e.) [Wave 1]	Mean WTP for the 1 in 1,000 risk reduction (s.e.) [Wave 2]
Level A = Minimum data cleaning	600.78 (25.88) $n = 625$	368.43 (29.89) $n = 298$
Level A + (Drop if flag1 = 1) (preferred cleaning procedure)	597.66 (25.75) $n = 611$	376.11 (29.51) $n = 286$
Level A + (Drop if flag4 = 1 and flag5 = 1)	597.44 (24.29) $n = 615$	373.21 (30.07) $n = 288$
Level A + (Drop if flag1 = 1 or flag2 = 1 or flag3 = 1 or flag4 = 1 or flag6 = 1)	593.05 (28.41) $n = 524$	386.16 (34.06) $n = 247$

Level A = drop respondents that answer the follow-up, open-ended question on WTP with an amount greater than the lowest bid previously received. For instance, a person who said he would not pay C$225, and would not pay $100, but when probed in an open-ended fashion said $300 is excluded from the sample.

Flag1 = 1 if respondent answers the first probability test incorrectly and shows preference for having the higher risk of death
Flag2 = 1 if respondent answers the probability test incorrectly and initially shows preference for having the higher risk of death, but changes preference selection when asked to confirm
Flag3 = 1 if respondent shows preference for having higher risk of death
Flag4 = 1 if respondent answers both probability tests incorrectly
Flag5 = 1 if respondent shows preference for higher risk of death even after follow up
Flag6 = 1 if respondent says that he does not understand probability well

has little effect on estimates of the VSL, which remains C$l.2 million for the 5 in 10,000 annual risk change and C$3.8 million for the 1 in 10,000 annual risk change.

Other, more stringent data cleaning criteria were also applied but, as shown in the other rows of Table 6, had very little effect on mean WTP estimates. This remains true when the sample is partitioned into two groups consisting of those respondents who were sure of their WTP responses, and of those respondents who did not feel confident, respectively. We note, however, that degree of confidence acts in opposite ways in the two samples. Those in wave 1 who were more certain of their responses are willing to pay more than those who were less certain, but the opposite result occurs for those in wave 2.[19]

5.5. Sensitivity of WTP *to age: 5 in 1,000 risk reduction*

Are the WTP amounts of older and sicker individuals different from those of younger and healthier individuals? To answer this question, we regressed WTP on age and health status variables, while controlling for factors such as income and education, probability comprehension and acceptance of the risk reduction scenario.

In Table 7 we present results of the "spike" model, estimated using our preferred cleaning procedure, wave 1 observations, and several covariates. The coefficients of the variables entered in the model should be interpreted as indicating the effect that a change in one variable has (all else unchanged), on conditional median WTP.[20] Table 7 shows the effect of age in three different model specifications. In specification 1, we impose a linear relationship between age and WTP, while in specification 2 the functional form is quadratic. Age-group dummies are entered in specification 3 to allow for greater flexibility in the relationship between age and WTP.

The table shows that WTP does not change in either a linear or quadratic fashion with age.[21] The results of specification 3 show that median WTP is similar across age groups up to about 70 years of age, but is lower for persons aged 70 years and older. One problem with this specification is that the estimated intercept, which captures respondents aged 70 years and older, is not significant, probably as a result of the low number of respondents in this age group.[22]

In Table 8, we calculate mean WTP for each age group by separating the data by the age of the respondent, and fitting to each sub-sample a tobit/spike model that includes only the intercept. Table 8 shows that mean WTP is about C$583 among the younger respondents in our sample (ages 40 to 49), rising to about C$657 for the next two age groups, with the oldest respondents (aged 70 and older) reporting mean WTP of about C$417. A Wald test shows that this latter figure is different from those for the immediately preceding age groups as well as that from the 40-to-49 year-olds.[23] When this estimation procedure is repeated using only two broad age categories—65 and older and younger than 65 years—the difference in WTP is smaller, but consistent with the preceding results: Average WTP is C$617 before age 65 and C$515 after age 64. The finer age breakdown, however, contains more information, and it is those results we emphasize: mean WTP after age 70 declines by about l/3 from its value between ages 50 and 70.[24]

Table 7. Internal validity of WTP: 5 in 1,000 risk change, wave 1, spike model, cleaned data ($N = 600$)

	Coefficient (t-statistic)		
Variable	Specification 1	Specification 2	Specification 3
Intercept	475.25	−1060.79	198.09
	(0.67)	(−1.08)	(0.67)
Age	−2.52	54.61	—
	(−0.84)	(1.53)	
Age squared	—	−0.51	—
		(−1.61)	
Ages 40 to 50	—	—	207.55*
			(2.04)
Ages 51 to 60	—	—	255.10**
			(2.38)
Ages 61 to 70	—	—	253.32*
			(2.29)
Male	−36.23	−31.94	−33.13
	(−0.60)	(−0.53)	(−0.55)
Bottom 25% of distribution of income in the sample (dummy)	−220.12**	−208.78**	−205.47**
	(−3.03)	(−2.88)	(−2.84)
Education (years of schooling)	−12.74	−13.73	−13.68
	(−1.26)	(−1.37)	(−1.36)
Chronic Illness Dummy	2.35	1.25	6.60
	(0.04)	(0.02)	(0.11)
Mental health score	4.56**	4.54**	4.62**
	(2.54)	(2.54)	(2.59)
Does not believe risk figures	−61.68	−57.85	−61.85
	(−0.79)	(−0.75)	(−0.80)
Other benefits	241.51**	237.61**	235.38**
	(3.88)	(3.83)	(3.80)
Did not consider if he can afford payments	−386.79**	−394.86**	−398.65**
	(−5.46)	(−5.57)	(−5.63)
Did not understand timing of payments	108.91	100.30	102.38
	(1.22)	(1.13)	(1.15)
Scale parameter	651.62**	649.24**	647.59**
	(22.44)	(22.43)	(22.90)
Log likelihood	−1308.09	−1306.80	−1305.18

Note: *indicates significance at the 5% level; **indicates significance at the 1% level. The data are subject to the base level of cleaning (level A, see Table 6). In addition, respondents who gave the wrong response to the probability test *and* wish to be the person with the higher probability of death in the probability choice question (Flag1 = 1) are excluded from this sample.

5.6. *Sensitivity of* WTP *to health status: 5 in 1,000 risk reduction*

We are also interested in the effect of variables measuring health status. One way of measuring health status is to use dummy variables for the five chronic conditions listed in Table 3—heart disease, high blood pressure, cancer, asthma and the complex bronchitis/emphysema/chronic cough.

Table 8. Mean WTP by age. Predictions based on cleaned data, wave 1

Age group	Predicted mean WTP (1999 C$)	Standard error of prediction (1999 C$)
Ages 40–50 ($N = 260$)	582.82	35.63
Ages 51–60 ($N = 165$)	656.88	58.37
Ages 61–70 ($N = 120$)	657.83	60.87
Ages 70 and over ($N = 66$)	417.99	54.30
Less than 65 ($N = 494$)	617.19	30.22
65 and over ($N = 117$)	514.69	52.15

There is some evidence that having cancer may raise WTP for a higher probability of survival. Table 9 shows that the coefficient of the cancer dummy is significant at the 8 percent level, and both its magnitude and significance level are robust to the inclusion of the mental health score and the use of age and age squared in lieu of age group dummies. The presence of cancer raises median WTP by C$258 to C$269 (depending on the specification), holding the values of the other regressors unchanged.

The representative person with cancer is older than the remainder of the sample (62.4 v. 53.7 years old), more likely to be a female, slightly less educated, and much *less* likely to question the risk figures presented to him or her in the questionnaire, but similar to respondents not affected by cancer in most other respects. The appropriate regression of Table 9 predicts that a person with these characteristics *and* cancer would be willing to pay, on average, C$837 (standard error C$157). A person with these characteristics but no cancer holds mean WTP values of C$568 (s.e. C$67). The Wald statistic for the null hypothesis of no difference between these two mean WTP is 2.48 (p-value $= 0.11$). We caution, however, that this result is based on only 26 respondents who reported having cancer, and that these respondents were all well enough to travel to the survey site.

The dummy variables for the presence of the other chronic illnesses are not significant predictors of WTP, either when entered in the regression individually, as shown in Table 9,[25] or when they are collapsed into a new dummy (CHRONIC), equal to 1 if any chronic illness is reported (Table 7).

Another way of measuring health status is to use scores from SF-36. The 36 questions on SF-36 are used to construct 8 indices, listed in the Appendix, Table A.1, that capture various dimensions of physical and mental health. Each index runs from 1 to 100 points, with higher values indicating better health.[26]

In practice, entering each of the 8 indices from the SF-36 questionnaire into the model (one at a time, to minimize collinearity problems, as shown in Table 9) produces significant results for only two indices, both measuring *mental*, not physical, well-being—the Mental Health score and the Role-Emotional Score. Both are positively related to WTP.[27]

Table 9. Coefficients and p-values of health variables in spike model for 5/1,000 risk reduction (each variable added separately)

	Coefficient (p-value)		
Variable	Specification (a)	Specification (b)	Specification (c)
Heart disease	16.37 (0.870)		
High blood pressure	87.53 (0.258)		
Cancer	269.72 (0.0867)		
Asthma	32.42 (0.737)		
Bronchitis, emphysema or chronic cough	−111.00 (0.198)		
Physical functioning score		−1.06 (0.491)	−1.30 (0.413)
Social functioning score		2.00 (0.120)	1.99 (0.162)
Role limitation due to physical problems score		1.21 (0.132)	1.20 (0.145)
Role limitation due to emotional problems score		1.72 (0.0495)	1.70 (0.0524)
Mental health score		4.45 (0.0122)	4.43 (0.0132)
Energy/vitality score		2.12 (0.157)	2.11 (0.172)
Bodily pain score		−.603 (0.621)	−.700 (0.571)
General health perceptions score		1.30 (0.365)	

Note: Specifications (a) and (b) include, in addition to the health variable listed, age group dummies, male, education, dummy for bottom 25% of the income distribution, Flag7, Flag11, Flag15, Flag16. Specification 3 also includes the Chronic Illness dummy (= 1 if any chronic illness reported. The data are subject to the base level of cleaning (level A, see Table 6). In addition, respondents who gave the wrong response to the probability test *and* wish to be the person with the highest probability of death in the probability choice question (Flag1 = 1) are excluded from this sample.

The Mental Health score is bipolar and measures an individual's mental health function and wellbeing. A median score indicates the absence of psychological distress while a score of 100 indicates that the respondent reports feeling peaceful, happy and calm all of the time. The role-emotional score is a related concept, and it stands to reason that the happier the respondent feels about life, the more he or she should be willing to pay for an increase in the chance of survival.

The results for the Mental Health score are robust to the inclusion of the Role-Emotional score (in specifications omitted from this paper) and/or the dummy for any chronic condition (in Table 7). The estimated coefficient on the Mental Health score implies that—regardless of the specification of the model—every additional point raises median WTP by about C$4.50.

5.7. *Other determinants of* WTP *for the 5 in 1,000 risk reduction*

Returning to Table 7, males tend to have lower WTP, but this effect is not statistically significant. Being in the lowest quartile of the income distribution significantly reduces median WTP by about C$205, while education does not affect WTP.

Table 7 also includes dummy variables that measure respondent acceptance of the scenario. It is comforting that people who did not believe baseline risks do not differ systematically from people who did. This suggests that people accepted the questionnaire's request to make decisions and answer questions "as if" the risk figures applied to them, even though the respondent may have privately disagreed with them.

Those persons who thought of "other benefits" of the product (presumably, improved quality of life) have median WTP values for mortality benefits that exceed *pure* mortality benefits by C$235. Finally, confirming earlier results, we find that WTP is remarkably lower for people who did not even consider whether they could afford the product—having already ruled out that they would purchase it.

5.8. *Smaller risk reductions*

We ran regressions similar to those of Table 7 for WTP for the 1 in 1,000 risk reduction (wave 2), but found that almost all of our regressors were insignificant. Only gender was significant at the 10% level with males willing to pay approximately C$150 less than females, while the coefficient of the dummy indicating that the respondent did not consider whether he or she could afford the product was significant at the 1% level. Once again, WTP was lower for people that did not consider whether they could afford the product.

There are two possible reasons why most of the regressors we examined are not statistically significant. The first possible reason is the relatively small sample size (n = 300), compared to the relatively large dispersion of WTP. The second possible reason is that a 1 in 1,000 risk reduction might be too small a risk reduction for people to be able to strongly relate their WTP to their socio-economic circumstances.

To discriminate between these two possible reasons, we experimented with taking several random sub-samples of 300 observations from the 611 data points of wave 1 (valuing the 5 in 1,000 risk reduction). In doing so, we asked the question: Had the sample size been smaller, and comparable to that of wave 2, would most of the regressors for the 5 in 1,000 risk reduction (wave 1) become insignificant? If so, the lack of significance of the coefficients of the 1 in 1,000 risk reduction might also be ascribed to insufficient sample size.

In practice, even with the reduced sample size, the coefficients of the regressors of WTP for the 5 in 1,000 risk reduction had signs and magnitudes similar to those for the full sample reported in Table 7, and those coefficients that were significant in Table 7 remained statistically significant at the conventional levels (even though the t-statistics were, of course, a bit smaller). This leads us to speculate that other reasons, perhaps the small magnitude of the risk reduction, lie behind the lack of significance of most coefficients in the equation for WTP for the 1 in 1,000 risk reduction.

6. Conclusions and implications for policy

Traditional estimates of the willingness-to-pay for mortality risk reductions are seldom conditioned on the age and health status of the respondent. Our survey is designed to provide credible estimates of such values that are applicable to mortality risk reductions associated with environmental policy. These values could also be used in benefit-cost analyses of other health and safety programs that primarily benefit older persons, many of whom may be in poor health.

We find that for 930 persons aged 40 to 75 in Hamilton, Ontario, WTP for risk reductions varies significantly with the size of the reduction. Mean WTP for an annual reduction in risk of death of 5 in 10,000 is about 1.6 times WTP for an annual risk reduction of 1 in 10,000. WTP is, therefore, sensitive to the size of the risk reduction, but not strictly proportional to it.

Our estimates of mean WTP translate into values of a statistical life of approximately C$3.8 million (1999 C$) for a 1 in 10,000 annual risk reduction and C$l.2 million for a 5 in 10,000 annual risk reduction, or U.S.$3.04 million and U.S.$0.96 million, respectively. These are 10 to 70 percent lower than Health Canada's age-adjusted VSL of C$4.3 million (1999 C$) and one-half (or less) the size of the U.S.$6 million (1999 US$) figure used by the USEPA.

The questionnaire also contains a variety of questions and tasks to check for understanding of probability and acceptance of our scenarios. While only a small fraction of the sample appears to fail the probability comprehension tests, up to a quarter of the sample have doubts about the abstract risk reduction scenario. Provided that we control for these individuals, regression models reveal that WTP is internally valid.

We find that WTP does not vary much by age, up to 65. Persons 40 to 49 years old do have slightly lower WTP than persons 50 years of age and older; however, mean WTP (C$657 for the 5 in 1,000 risk change) remains approximately constant age until about age 70, decreasing by about one-third thereafter. This latter WTP (C$417) is probably the most relevant one for use in valuing most of the lives saved by air pollution reductions.

Regardless of the measure of physical health status used, WTP (with one exception), does not vary appreciably with physical health status either—an important result for environmental policy, since older people and people with chronic conditions are often the beneficiaries of improvements in environmental quality. We do, however, find that individuals with cancer are WTP over 60% more for a mortality risk reduction than their counterparts without cancer, and that individuals in better mental health have a larger WTP than those scoring lower on tests of their mental health.

These results stand in sharp contrast to the way in which age and health status are treated in evaluating medical interventions. We believe the comparison is relevant, since it is sometimes suggested that a similar approach be used in benefit-cost analyses of health and safety regulations (Food and Drug Administration, 1999). The standard approach in the medical literature is to measure life-saving benefits in terms of Quality Adjusted Life Years Saved (QALYs). This assumes that the value of lives saved is strictly proportional to remaining life expectancy, and that the value of saving a life-year is less for a person with a chronic disease, such as chronic bronchitis, than for a healthy person, with the exact equivalence determined by QALY weights. Our results do not support either of these assumptions. There is no evidence that the VSL should be equally apportioned over remaining life expectancy, or that the VSL is systematically lower for persons with chronic illness.

Appendix

Table A.1. Additional index scores from SF-36 questions

	Mean (std. dev.)		
Variable name	Wave 1 ($N = 630$)	Wave 2 ($N = 300$)	Total sample ($N = 930$)
Role-physical score (1–100) (Measures the extent of disability in everyday activities due to physical problems. Low score indicates problems with work or other daily activities resulting from physical health; High score indicates no problems with work or other daily activities as a result of physical health.)	74.6 (38.0)	77.7 (35.0)	75.6 (37.1)
Bodily pain score (1–100) (Measures the severity of bodily pain and resulting limitations in activities. Low score indicates very severe and extremely limiting pain; High score indicates no pain or limitations due to pain.)	67.9 (24.5)	70.3 (20.6)	68.7 (24.9)
General Health score (1–100) (Bipolar scale measuring respondents' perceived general health. Low score indicates respondent evaluates personal health as poor and believes it is likely to get worse; High score indicates the respondent evaluates personal health as excellent. Midrange score indicates no unfavorable evaluations of health.)	69.8 (21.0)	70.7 (20.5)	70.1 (20.9)
Vitality score (1–100) (Bipolar scale measuring energy level and fatigue; Mid-range score indicates that the respondent does not report feeling tired or worn out; Score = 100 indicates that in addition, respondent feels full of pep and energy all of the time.)	62.7 (20.1)	63.4 (20.0)	62.9 (20.1)

(Continued)

Table A.1. (Continued)

Variable name	Mean (std. dev.) Wave 1 (N = 630)	Wave 2 (N = 300)	Total sample (N = 930)
Social functioning score (1–100) (Measures the impact of either physical or emotional problems on the quantity and quality of social activities. Low score indicates extreme and frequent interference with normal social activities due to physical or emotional problems; High score indicates respondent performs normal social activities without interference due to physical or emotional problems.)	83.6 (22.9)	85.4 (21.5)	84.2 (22.4)
Role-emotional score (1–100) (Measures the extent of disability in everyday activities due to emotional problems. Low score indicates problems with work or other daily activities resulting from emotional problems; High score indicates no problems with work or other daily activities as a result of emotional problems.)	79.9(33.8)	83.3 (30.8)	81.0 (32.9)

Table A.2. Identifying reasons for WTP equal to zero: Probit model
Wave 1 ($N = 614$)
Zero WTP for 5/1,000 risk reduction

Variable	Estimate	*T* statistic
Intercept	−0.9603	−0.33
Baseline risk		
Male	−0.0224	−0.17
Age	0.0521	0.47
Age squared	−0.0005	−0.51
Individual characteristics:		
Income (thou. C$)	−0.0018	−0.96
Has chronic illness	−0.1144	−0.80
SF-36 mental health score	−0.0095*	−2.43
Understanding of probability/acceptance of scenario		
Failed probability test and preference	−0.6173	−1.09
Did not believe risks apply to him or her	0.1552	0.95
Did not consider whether (s)he could afford payment	0.7713*	5.35
Did not understand timing of payment	−0.3337	−1.51

t- statistics in parentheses; *significant at 1%.

Acknowledgments

The authors would like to thank David Stieb our project officer at Health Canada, Paul De Civita, also of Health Canada, the U.S. Environmental Protection Agency, Office of Policy for funding and many fine suggestions. Survey development was also supported by an

EPA/NSF grant and from a project initiated by MITI, an agency of the Japanese government, as well as RFF. We would also like to thank Robert Belli (University of Michigan) for his help in early stages of survey development, as well as Richard Carson, Lauraine Chestnut, William Desvousges and Magnus Johannesson, who reviewed an early version of the survey. Finally, we wish to thank the many participants of workshops and conferences where we presented earlier draft papers for their useful and insightful comments. The findings, interpretations and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of Health Canada, the USEPA or the World Bank, its Executive Directors or the countries they represent.

Notes

1. Using purchasing power parity of \$1.25 C to \$1.0 U.S. The currency exchange rate at the time of the study was \$1.5 C to \$1.0 U.S.
2. This assumption is implicit in time-series models in which deaths on day t are assumed to be an exponential function of air pollution on day $t - s$, weather and other variables. It is also embodied in the prospective cohort study of Pope et al. (1995), which assumes that the impact of air pollution is proportional to the probability of dying at each age (given that one survives to that age).
3. Cancer is the health endpoint most often quantified in connection with hazardous waste sites, pesticide regulations and drinking water standards.
4. In 1996, 71 percent of all cancer deaths in the U.S. occurred among residents aged 65 years and over (U.S. Census Bureau, 1999).
5. To illustrate this calculation, suppose that the value of a statistical life based on compensating wage differentials is \$5 million, and that the average age of people receiving this compensation is 40. If remaining life expectancy at age 40 is 35 years and the interest rate is zero, then the value per life year saved is approximately \$140,000. If, however, the interest rate is 5 percent, then *discounted* remaining life expectancy is only 16 years, and the value per life-year saved rises to approximately \$300,000.
6. Moore and Viscusi (1988) have used labor market data to infer the rate at which workers discount future utility of consumption; however, their models make very specific functional form assumptions in order to infer a discount rate from a single cross section of data.
7. We restricted our sample to individuals of ages 40–75 for two reasons. First, we realized during the initial questionnaire development that it was very difficult to get younger subjects to focus on their (generally very low) risk of dying over the next 10 years. Hence it was difficult to design a questionnaire that would apply to persons 18–75. Second, in the air pollution context (and for other environmental policies) the policy debate is about whether estimates of VSL should be adjusted for older people and people in poor health. Restricting our sample to older age groups allows us to address this question.
8. A pure altruist cares about his neighbor's health but respects his preferences. If the neighbor must pay for the health or safety program, the pure altruist realizes that the program will impose costs as well as benefits on his neighbor, and respects the rate at which the neighbor is willing to substitute income for health. This rate is already captured in the neighbor's private WTP for the program. If, however, the altruist exhibits paternalistic or safety-focused altruism, caring only for his neighbor's health and safety and not about his wealth, then it would be correct to include the altruist's WTP for the recipient's health/safety when determining the socially optimal level of health/safety (Jones-Lee, 1991).
9. Our decision to provide qualitative information about the cost of risk-reducing behaviors was motivated by three considerations. First, we did not want to provide dollar figures that might anchor the respondent's answers to the WTP questions later in the survey. Second, the labels of 'inexpensive,' 'moderate' etc. were intended to describe risk-reducing behaviors in a relativistic sense: For example, exercise as a preventive measure would be 'inexpensive' *relative to* annual screening for certain types of cancer, and so forth.

Third, providing qualitative rather than quantitative information helped reduce the amount of quantitative information that the respondent was asked to process in the survey.

10. Respondents in both waves over age 60 were also asked what they would pay over the next ten years for a 5 in 1,000 risk reduction (5 in 10,000 risk reduction annually, over ten years), beginning at age 70. The analysis of these responses appears in Krupnick et al. (2000).
11. The response rate is calculated as the number of respondents successfully completing the study (930) divided by the number of eligible contacts (3,591).
12. Descriptive statistics for other index scores, and definitions for all SF-36 index scores, are reported in the Appendix, Table A.1.
13. Those respondents who received the 1 in 1,000 risk reduction first (wave 2) exhibit higher WTP than those respondents who received it as the second risk reduction in the questionnaire (wave 1). Since Tables 4, 5 and 6 show clearly that the two waves of respondents are very similar in terms of demographics, probability comprehension and acceptance of the life-extending product, we believe that the effect noted above is questionnaire-induced.
14. The percentage of "yes" responses declines as the bid increases for the l/1,000 risk reduction, except at the initial bid amount of C$1,100, where it is virtually the same as that for the initial bid amount of C$750 (19.67% vs. 19.32%, respectively).
15. The payment questions actually offered respondents three response options: "yes," "no" and "not sure." The percentage of "not sure" was not very high, and increased monotonically with the bid value. We fit multinomial logit models relating the choice of a response category to the bid level and to individual characteristics and answers to the debrief questions, finding that the coefficients of the bid and most variables were statistically indistinguishable across the "no" and "not sure" response categories (see Krupnick et al., 2000). As in Carson et al. (1998), we interpret this as evidence that the "not sure" responses can be interpreted and statistically modeled as if they were "no" responses, which we do in this paper.
16. WTP passes the scope test, but fails to be strictly proportional to the size of the risk change, even when one focuses on *internal* comparisons, examining mean WTP for the two risk reductions received by each respondent. Specifically, respondents in wave 1 hold mean WTP values of C$598 (standard error C$27) and C$240 (s.e. C$16) for the 5 in 1000 and 1 in 1000 risk reduction, respectively. These figures are statistically different (Wald statistic 126.22; p-value < 0.0001), but not strictly proportional to the risk change (the ratio is 2.48, with a Wald statistic of 49.56, which soundly rejects the null hypothesis of proportionality). Similarly, respondents in wave 2 imply mean WTP values of C$496 (s.e. C$35) and $376 (s.e. C$30) for the 5 in 1000 and 1 in 1000 risk reductions. While the two mean WTP amounts are statistically different (Wald statistic 6.80, which rejects the null of no difference at the 1% level), their ratio is only 1.31 and the null hypothesis of proportionality is rejected at all conventional significance levels (Wald statistic 80.91, p-value < 0.0001).
17. The VSLs are calculated assuming that the risk change is evenly distributed over the 10-year period. Because the payment and the risk change are discounted over the same period of time using the same discount rate, the choice of discount rate in this scenario is irrelevant.
18. Alternative ways of calculating mean WTP, and hence VSL, are possible. To get distribution-free and conservative estimates of mean WTP, we applied the lower-bound Turnbull estimation technique described in Carson et al. (1994). This approach utilizes only the responses to the initial payment questions, ignoring the responses to the follow-up questions, and is thus safe from undesirable response effects sometimes observed in the presence of follow-up questions (Herriges and Shogren, 1996; Alberini, Kanninen and Carson, 1997). The approach produces estimates of mean WTP equal to $361 (5 in 1000 risk reduction, wave 1) and $272 (1 in 1000 risk reduction, wave 2), which are statistically different from each other, but fail the proportionality test. The resulting VSLs are $722,000 and $2,272,000, respectively. For "generous" estimates of WTP, we experimented with ignoring respondents' final announcements that they were not willing to pay anything at all for the risk reduction, assuming instead that the lower and upper bounds of the interval around their WTP amounts are zero and the lowest bid amount they were offered. We then fitted a Weibull distribution to the mix of continuous and interval data derived in this fashion, and obtained estimates of mean WTP equal to $726 (5 in 1000 risk reduction, wave 1) and $450 (1 in 1000

risk reduction, wave 2). Once again, these figures are statistically different, but the former is less than five times the latter. The implied VSLs are $1,452,000 and $4,496,000 ($C). We conclude that, regardless of the approach used, WTP is sensitive to the size of the risk reduction, but not strictly proportional to it, and VSL is less than the official figures used by either Health Canada or the US EPA.

19. While the ratio between mean WTP for the 5 in 1000 risk reduction and mean WTP for the 1 in 1000 risk reduction remains less than 5, it is slightly higher when the mean WTP figure are computed for those respondents who feel more confident about their responses to the payment questions, as in other studies surveyed by Hammitt and Graham (1999).
20. Mean WTP is a non-linear function of the regression coefficients. For respondent i, mean WTP is $\Phi(\mathbf{x}_i, \beta/\sigma) \cdot [\mathbf{x}_i, \beta + \sigma(\phi(\mathbf{x}_i, \beta/\sigma)/\Phi(\mathbf{x}_i, \beta/\sigma)]$, $\mathbf{x}$ is the vector of regressors and β is the vector of regression coefficients.
21. In the quadratic specifications, the coefficients of age and age squared are individually insignificant. In addition, a likelihood ratio test fails to reject the null hypothesis that they are *jointly* equal to zero (likelihood ratio statistic 2.40, p-value 0.30).
22. When the tobit/spike regression models of Table 7 are repeated for different data cleaning criteria, the results are generally similar. The only difference we noticed is that, when a relatively stringent cleaning criterion is used (excluding respondents who admit poor understanding of the concept of probability, and, as before, fail the probability test and choice questions), the coefficient of age and age squared become strongly significant, providing evidence of an inverted U-shaped function relating WTP to age. The peak WTP value is reached at age 57, but the curvature of the inverted-U curve is not very pronounced, confirming that the impact of age is modest. In practice, we feel that the results from the specification that is quadratic in age should be interpreted with care, because of the very high linear dependence between age and age squared over the range of ages of our respondents.
23. The Wald statistics are equal to 8.64, 8.97 and 6.44 respectively, and thus fall in the 1% rejection region for the chi square with one degree of freedom.
24. The effect of age on WTP is not altered by the inclusion of dummies indicating how the respondent interpreted the risk reduction scenario (see Section 5.7), as the latter do not appear to depend age. The age effects are virtually unchanged when such dummies are omitted from the right-hand side of the WTP model.
25. Table 9 reports an estimated coefficient of C$16 for heart disease, $87 for high blood pressure, $34 for asthma and $111 for bronchitis, emphysema or chronic cough, but none of these coefficients is significant at conventional levels.
26. Pair-wise correlations among the 8 indices are all positive and significant ($p < .0001$), with correlation coefficients ranging from 0.3 to 0.6. Thirty-three out of 40 pair-wise correlations between the SF-36 scores and the chronic illness dummies are negative and significant at the .05 level or better.
27. Using quadratic functions of the physical health scores, or creating splines, did not change the nature of the results.

References

Alberini, Anna, Barbara Kanninen, and Richard T. Carson. (1997). “Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data,” *Land Economics* 73(3), 309–324.

Bousquet, J., et al. (1994). “Quality of Life in Asthma. I. Internal Consistency and Validity of the SF-36 Questionnaire,” *American Journal of Respiratory and Critical Care Medicine* 149, 371–375.

Canada-Wide Standards Development Committee for PM and Ozone. (1999). “Discussion Paper on Particulate Matter and Ozone,” Canada-Wide Standards for Consultation, National Multi-Stakeholder Consultation Workshop, Calgary, Alberta, May 26–28.

Carson, Richard T., et al. (1998). “Referendum Design and Contingent Valuation: The NOAA Panel’s No-Vote Recommendation,” *Review of Economics and Statistics* 80(3), 484–487.

Carson, Richard T., et al. (1994). “Prospective Interim Lost Use value due to DDT and PCB Contamination in the Southern California Bight,” Report prepared under NOAA contract No. DGNC-1-00007.

Cousineau, Jean-Michel, Robert Lacroix, and Anne-Marie Girard. (1992). "Occupational Hazard and Wage Compensating Differentials," *The Review of Economics and Statistics* 74(1), 166–169.

EOP Group, Inc. (1997). *Life-Year Analysis of Premature Mortality Benefits in the December 1996 Particulate Matter Proposed NAAQS.* February.

Fahrmeir, Ludwig, and Gerhard Tutz. (1994). *Multivariate Statistical Modeling Based on Generalized Linear Models.* New York: Springer-Verlag.

Food and Drug Administration—HHS. (1999). "Preliminary Regulatory Impact Analysis and Initial Regulatory Flexibility Analysis of the Proposed Rule to Require Refrigeration of Shell Eggs at Retail and Safe-handling Labels," *Federal Register* 64(128), July 6.

Gerking, Shelby, Menno de Haan, and William Schulze. (1988). "The Marginal Value of Job Safety: A Contingent Valuation Study," *Journal of Risk and Uncertainty* 1(2), 185–200.

Hammitt, James K. and John D. Graham. (1999). "Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability?" *Journal of Risk and Uncertainty* 8, 33–62.

Herriges, Joseph A. and Jason F. Shogren. (1996). "Starting Point Bias in Dichotomous Choice Valuation with Follow-Up Questioning," *Journal of Environmental Economics and Management* 30, 112–131.

Johansson, Per-Olov. (1994). "Altruism and the Value of Statistical Life: Empirical Implications," *Journal of Health Economics* 13, 111–118.

Jones-Lee, Michael W. (1991). "Altruism and the Value of Other People's Safety," *Journal of Risk and Uncertainty* 4, 213–219.

Jones-Lee, Michael W., M. Hammerton, and P. R. Phillips. (1985). "The Value of Safety: Results of a National Sample Survey," *Economic Journal* 95, 49–72.

Kniesner, Thomas J. and John D. Leeth. (1991). "Compensating Wage Differentials for Fatal Injury Risk in Australia, Japan and the United States," *Journal of Risk and Uncertainty* 4(1), 75–90.

Kriström, Bengt. (1997). "Spike Models in Contingent Valuation," *American Journal of Agricultural Economics* 79, 1013–1023.

Krupnick Alan, et al. (2000). "Age, Health, and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents," RFF Discussion Paper QE00-37, Washington, DC, September.

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

Miller, Ted R. and J. Guria. (1991). *The Value of Statistical Life in New Zealand*. Wellington, NZ: Land Transport Division, New Zealand Ministry of Transport.

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

Pope, C. Arden, et al. (1995). "Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults," *American Journal of Respiratory Critical Care Medicine* 151, 669–674.

Schwartz, Joel. (1991). "Particulate Pollution and Daily Mortality in Detroit," *Environmental Research* 56, 204–213.

Schwartz, Joel and Douglas W. Dockery. (1992). "Increased Mortality in Philadelphia Associated with Daily Air Pollution Concentrations," *American Review of Respitory Disease* 145, 600–604.

Schwartz, Joel. (1993). "Air Pollution and Daily Mortality in Birmingham, Alabama," *American Journal of Epidemiology* 137(10), 1136–1147.

Smith, V. Kerry and William H. Desvousges. (1987). "An Empirical Analysis of the Economic Value of Risk Changes," *Journal of Political Economy* 95, 89–113.

Stieb, David. (2000). Personal communication, Health Canada.

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

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

Viscusi, W. Kip. (1992). *Fatal Tradeoffs. Public and Private Responsibilities for Risk*. Oxford: Oxford University Press.

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

Ware, John E., Jr., Mark Kosinski and Susan Keller. (1997). *SF-36 Physical and Mental Health Summary Scales A User's Manual.* Lincoln: RI: Quality Metric.