

**“Peeling back the onion:  
Using latent class analysis to uncover heterogeneous  
responses to stated preference surveys”**

James K. Hammitt and Daniel Herrera-Araujo

Peeling back the onion:  
Using latent class analysis to uncover heterogeneous responses to  
stated preference surveys

James K. Hammitt\*  
Daniel Herrera-Araujo†

January 28, 2017

**Abstract**

We develop validity tests for application to stated-preference estimates of WTP to reduce mortality risk, i.e., value per statistical life (VSL), and apply these to data obtained by surveying a representative sample of French adults over the internet. These tests (WTP nearly proportional to risk reduction, insensitive to small differences in baseline risk, increasing in income, and consistent with budget constraints) are satisfied by a conventional single-regression analysis of our data. Using latent class analysis (LCA), we identify important differences between respondents in their consistency with the validity tests and control for much of this heterogeneity. Estimates of VSL from the latent class that is consistent with the validity tests are smaller than estimates from the standard analysis. We estimate mean VSL for adults of about 2 million € and for children (based on parents' WTP) of about 6 million €

**JEL Codes:** D03, D61, D64, I18, Q18, Q51.

**Keywords:** Value per Statistical Life, Latent Class Analysis, Paradata, Scope sensitivity, Income elasticity, Pesticides, Children.

---

\*Corresponding author. Center for Risk Analysis, Harvard University, 718 Huntington Ave., Boston, MA 02115, USA, and Toulouse School of Economics, Université Toulouse Capitole, 21 allée de Brienne, 31000 Toulouse, France. Phone: +1.617.432.4343, E-mail: jkh@harvard.edu

†Paris School of Economics, Hospinnomics, Paris, France, E-mail: daniel.herrera@psemail.eu

# 1 Introduction

The value of reductions in mortality risk, conventionally measured using the value per statistical life (VSL), accounts for the lion's share of quantified benefits of many environmental, health, and safety regulations (Hammitt & Robinson 2011).<sup>1</sup> The appropriate VSL to be applied in these analyses has received substantial attention. As markets for mortality-risk reduction do not exist, researchers have used non-market valuation techniques, including revealed preference (RP) and stated preference (SP) methods. RP methods infer individuals' preferences from their behavior in contexts that affect their mortality risks. RP has most often been used with labor market data, assessing the compensating wage differentials workers receive for occupational fatality risk (Viscusi & Aldy 2003). SP methods are based on respondents' responses to hypothetical choices. Two methods dominate the SP literature: contingent valuation methods (CVM, Mitchell and Carson 1989) and discrete choice experiments (Bateman et al. 2002).<sup>2</sup>

SP methods have been more controversial than RP; Hausman (2012) entitles his critique of the CVM, "From dubious to hopeless." Despite criticism, the CVM remains one of the principal sources for estimating the rate of substitution between wealth and small changes in health risks. Diamond et al. (1994) ask "Is some number better than no number?" and argue that if CVM does not accurately elicit preferences, the answer is no. Beshears et al. (2008) distinguish between revealed preferences that "rationalize an economic agent's observed actions" and normative preferences that "represent the agent's actual interests." On the assumption that policy evaluation should be informed by normative preferences, the critical question is under what conditions can SP estimates be interpreted as informative about normative preferences? A partial answer is that estimates must be consistent with standard models of tradeoffs; e.g., WTP should be sensitive to scope (larger for a greater benefit).

To assess the validity of SP estimates of willingness to pay (WTP) for small reductions in mortality risk, we derive validity tests from an expected utility model and use these to evaluate responses from a contingent valuation survey fielded to a panel representative of the French adult population. Our validity tests include: (1) WTP should increase with the magnitude of the risk reduction and (for the small risk reductions in question) should be nearly proportional to the risk reduction (Corso et al., 2001; Alolayan et al., 2015); (2) WTP should not vary significantly with small changes in baseline mortality risk; (3) WTP

---

<sup>1</sup>A retrospective analysis of the Clean Air Act finds that mortality risks account for 95 percent of the present value of monetized benefits from 1970 to 1990 (EPA 1997). Hence different values of VSL may substantially change the set of alternative policies that are judged cost-beneficial.

<sup>2</sup>VSL is defined as an individual's marginal rate of substitution between wealth and mortality risk in a specified period, such as the current year (Hammitt 2000). In the United Kingdom, benefits of policies to reduce mortality risk are evaluated using the Value of Preventing a Statistical Fatality (VPF), "conventionally defined as the aggregate willingness to pay by a large, representative sample of individuals for small reductions in the risk of death which, taken across the group concerned, will reduce the expected number of fatalities during a forthcoming period by one" (McDonald et al. 2016, p. 138). The VPF "incorporates the 'extra' value placed on relatives and friends, and any further value placed by society on avoiding the premature death of individuals" and in practice may also include "gross lost output, medical and ambulance costs" (HM Treasury 2011, p. 62).

should not decrease with wealth; and (4) WTP should not exceed the respondent's income. We examine several alternative estimates of VSL, including a novel "scope-revealing" VSL that equals zero when responses are completely insensitive to scope.

We explore the effects of heterogeneity among responses using latent class analysis (LCA). LCA endogenously creates classes composed of relatively homogeneous responses; each class is a weighted average of respondents and each respondent has a positive probability of membership in each class. LCA can be interpreted as a smoother method for weighting respondents than the common approach in which respondents who fail tests of comprehension, provide protest or extreme responses, or violate other criteria are excluded (given zero weight) and all other respondents are given equal weights.

Our key finding is that LCA identifies substantial heterogeneity among responses. Analyzing our data using a conventional regression approach yields results that satisfy our validity criteria. However, peeling back the onion with LCA reveals substantial heterogeneity in the extent to which the validity criteria are satisfied. Overall, 58 percent of responses are associated with latent classes that are inconsistent with our criteria. In one class, including 29 percent of responses, estimated WTP greatly exceeds income, likely violating respondents' budget constraints. In another class, accounting for another 29 percent of responses, estimated WTP significantly decreases with baseline risk, violating the criterion of insensitivity to baseline risk. The remaining class, associated with 42 percent of responses, satisfies our validity criteria. We suggest that estimates of VSL for this class may be interpreted as valid estimates for the population.

An additional benefit of controlling for heterogeneity using LCA is that the residual variance around the regression function is substantially reduced. Consistent with much of the previous literature, we analyze the logarithm of WTP as a function of demographic and other covariates; retransforming predicted log WTP to predicted WTP in euros can yield a large difference between the predicted mean and median.<sup>3</sup>

Our survey instrument is adapted from one administered to a representative internet panel in the United States by Hammitt and Haninger (2010). We elicit estimates of WTP to reduce mortality risks to identified individuals: the respondent him or herself, his or her child, another adult living in the household, and jointly to everyone living in the household. Hypothetical mortality risks are associated with pesticide residues on food, and risk reductions are obtained by purchasing an alternative food produced following a hypothetical, "pesticide security system" warranted by the state. The alternative food is described as produced using pesticides that are safer to humans than the pesticides used in producing conventional food (i.e., the alternative food reduces risk to human consumers but does not present environmental, worker-safety, or other benefits). Risks are described as a function of baseline risk of illness (with the conventional food type), risk reduction (with the alternative food type), disease type (cancer, non-cancer), affected organ (brain, bladder, liver, lymphocytes), latency period (1, 10, 20 years) and symptom description

---

<sup>3</sup>The predicted mean is larger than the predicted median by a factor of  $\exp(\frac{\sigma^2}{2})$ , where  $\sigma^2$  is the residual variance.

(provided or not). These characteristics are randomly varied across respondents and the person at risk (e.g., self, child) using a full factorial design. Estimates are obtained using a representative French internet panel. A total of 1000 respondents completed the survey.

Respondents value risk reductions to several members of their household ("targets"). The order of the targets is randomly assigned. We find that respondents' answers to the valuation questions about the first target exhibit greater scope sensitivity than answers concerning the second and third targets. This finding is robust across model specifications, even after controlling for observed and unobserved heterogeneity. As identification relies on between-respondent variation, it does not reflect a tendency of respondents toward arbitrary coherence, in which responses to valuation questions are motivated by consistency with responses to previous questions (Ariely et al. 2003, Goldberg & Rosen 2007).

Our analysis provides new estimates of VSL for the French population and describes how VSL varies with characteristics of the disease and the affected individual. Despite its importance to policy evaluation, there are few estimates of VSL in France.<sup>4</sup> Most of the French studies violate criteria for economic validity.<sup>5</sup> We propose VSL estimates for France that satisfy validity criteria. Our estimates range from 2 to 6 million € for adults and 6 to 7 million € for children. Consistent with some recent literature, we find no evidence that WTP is larger to reduce risk of cancer than other fatal diseases (Hammit & Haninger 2010; Chestnut et al. 2012; Cameron & Deshazo 2010).<sup>6</sup> LCA proves useful to understanding heterogeneity among respondents. Our preferred estimates are derived from one of the latent classes, and hence on an unequal weighting of respondents. This creates a possibility of bias compared with estimates based on equal weighting of respondents, but the distribution of demographic characteristics for the weighted sample is similar to that for the unweighted sample.

The paper is organized as follows: Section II provides theoretical background and derives predictions about WTP that are used as criteria to evaluate the extent to which the CV estimates can be interpreted within a standard economic model; it also presents an alternative, "scope-revealing" measure of VSL. Section III provides details on the survey design. Section IV describes the econometric models. Section V presents the main results and Section VI reports robustness checks. Section VII discusses the results and concludes.

---

<sup>4</sup>In 2013 the French administration updated its guidelines for project evaluation with the Rapport Quinet. The VSL it endorsed was extracted from a 2012 OECD meta-analysis, which contained only four studies eliciting VSL in France (Lindhjem et al. 2012). Three estimate monetary values for a risk reduction associated with pollution and one with transportation. None value risk reductions for cancer or other degenerative diseases.

<sup>5</sup>For example, using the same questionnaire as Alberini et al. (2006), Desaignes et al. (2007) estimated the value of life expectancy gain due to a reduction of air pollution in France. They report a large embedding effect (Kahneman et al. 1992); i.e., the ratio of WTP to reduce risk by 5/1000 to WTP to reduce risk by 1/1000 is 1.6; theory suggests that this ratio should be close to 5. The estimated VSL is 4.1 million €.

<sup>6</sup>In contrast, there is evidence that WTP to reduce risk of fatal cancer exceeds WTP to reduce risk of fatal motor-vehicle crashes, perhaps because of longer or more severe morbidity prior to death from cancer and greater dread of cancer (Sunstein 1997). Van Houtven et al. (2011), Viscusi et al. (2014), and McDonald et al. (2016) find evidence of a positive cancer premium compared with traffic fatality.

## 2 Model

We model consumer WTP to reduce mortality risk using a conventional single-period, state-dependent expected utility framework (Drèze 1962, Jones-Lee 1974, Weinstein et al. 1980). Let utility depend on wealth ( $w$ ) and on whether one survives the current period,  $u_j = u_j(w)$ , where  $j = A, D$  denote the two states, alive or dead, respectively. The utility of wealth conditional on death is associated with bequest motives. We adopt the standard assumptions  $u_A > u_D$ ,  $u'_A > u'_D \geq 0$ ,  $u''_A \leq 0$ ,  $u''_D \leq 0$ .

Letting  $\pi$  denote the probability of survival, expected utility is given by  $E(\mathcal{U}) = \pi u_A(w) + (1 - \pi) u_D(w)$ . Willingness to pay to reduce the risk by the amount,  $e$ , denoted  $P(e, w, \pi)$ , is defined by:<sup>7</sup>

$$(\pi + e) u_A(w - P(e, w, \pi)) + (1 - \pi - e) u_D(w - P(e, w, \pi)) = \pi u_A(w) + (1 - \pi) u_D(w). \quad (1)$$

Note that when  $e = 0$  then  $P(e, w, \pi) = 0$ .

The marginal rate of substitution between WTP,  $P$ , and risk reduction,  $e$ , is:

$$\frac{\partial P(e, w, \pi)}{\partial e} = \frac{u_A(w - P(e, w, \pi)) - u_D(w - P(e, w, \pi))}{(\pi + e) u'_A(w - P(e, w, \pi)) + (1 - \pi - e) u'_D(w - P(e, w, \pi))} > 0. \quad (2)$$

We define the value per statistical life, VSL, as the slope of the WTP function evaluated at zero risk reduction:

$$VSL = \frac{\partial P(0, w, \pi)}{\partial e} \equiv \frac{\partial P_0}{\partial e}. \quad (3)$$

Let  $\eta_e^P$ ,  $\eta_w^P$ , and  $\eta_{1-\pi}^P$ , denote the elasticity of willingness to pay  $P(e, w, \pi)$  with respect to the risk reduction  $e$ , wealth  $w$ , and baseline risk,  $1-\pi$ , respectively. Moreover denote by  $\eta_w^{VSL}$  the wealth elasticity of VSL. The following results hold:

$$\lim_{e \rightarrow 0} \eta_e^P = 1, \quad (4)$$

$$\eta_w^{VSL} = \lim_{e \rightarrow 0} \eta_w^P > 0, \quad (5)$$

$$\frac{1 - \pi}{\pi} \geq \lim_{e \rightarrow 0} \eta_{1-\pi}^P = \frac{1 - \pi}{\pi + \frac{u'_A(w)}{u'_A(w) - u'_D(w)} - 1} > 0, \quad (6)$$

$$P(e, w, \pi) < w. \quad (7)$$

Hence, for any utility function satisfying our assumptions and sufficiently small risk reduction, a small proportional increase of the risk reduction increases WTP by the same

<sup>7</sup>We assume that both  $e$  and  $\pi$  are exogenous to the individual.

<sup>8</sup>If we assume the utility of bequest  $u_D = 0$  we can re-express equation (1) as:  $P(e, w, \pi) = w - u_A^{-1}\left(\frac{\pi}{\pi+e}u_A(w)\right)$ . Here, it is clear that  $P(e, w, \pi) < w$ .

proportion, an increase of wealth increases WTP, and a small increase of baseline risk has virtually no effect on WTP.<sup>9 10</sup>

Under the standard expected-utility model, equations (4), (5), (6) and (7) provide powerful, yet simple, testable implications, which are key in assessing whether a contingent valuation survey is valid.<sup>11</sup> Table 1 summarizes the empirical tests to be performed.

Table 1: Validity theoretical tests summary

Characteristics	Criterion	Name of the test
Risk reduction	$\eta_e^P \approx 1$	$e$ -test
Baseline risk	$\eta_{1-\pi}^P \approx 0$	$\pi$ -test
Income	$\eta_w^P \geq 0$	$w$ -test
Budget	$w > P(e, w, \pi)$	$P$ -test

As we elicit WTP for household-level risk reductions, we provide two additional tests corresponding to our specific context. Each can be regarded as a test of scope sensitivity: (1) WTP for a risk reduction affecting everyone in a household cannot be less than the minimum WTP to reduce risk to any individual living in the household; and (2) when households are composed of a single individual, the differences between WTP to reduce risk to the household and to the individual alone should be zero.

Suppose a consumer is offered a risk reduction of size,  $e_1$ . We know that he or she is willing to pay an amount,  $P(e_1, w, \pi)$ . Suppose she is offered an alternative risk reduction of size  $e_2$ , for which she is willing to pay an amount  $P(e_2, w, \pi)$ . VSL is typically estimated as WTP divided by the risk reduction, i.e.,

$$VSL \approx VSL_{e_1} \equiv \frac{P(e_1, w, \pi)}{e_1}, \quad (8)$$

$$VSL \approx VSL_{e_2} \equiv \frac{P(e_2, w, \pi)}{e_2}. \quad (9)$$

By equation (2), if  $e_2 > e_1$  it must be the case that  $P(e_2, w, \pi) > P(e_1, w, \pi)$ . Then,

<sup>9</sup>Derivations of the three functional relationships are in the appendix.

<sup>10</sup>Under our model, risk reduction is exogenous. Liu et al. (2006) examine a model in which risk reduction depends on two inputs: the individual's private expenditure and an external expenditure. In their model, the sign of the effect of baseline risk on marginal WTP is not determined. If an increase of the external expenditure does not make the private expenditure more productive, then marginal WTP can decrease with baseline risk. The signs of the effects of risk reduction and income on marginal WTP are not affected.

<sup>11</sup>Hammitt (2000) considers the implications of alternative, non-expected-utility models of decision making under uncertainty. He argues that the sole requirement to satisfy near-proportionality is local linearity in probabilities (Machina 1992). If, for example individuals are willing to pay for a risk reduction only if the reduction exceeds a threshold that they consider meaningful, near proportionality need not hold.

by the mean value theorem, there exists a risk reduction  $e_c$  such that:

$$\frac{\partial P(e_c, w, \pi)}{\partial e} = \frac{P(e_2, w, \pi) - P(e_1, w, \pi)}{e_2 - e_1}. \quad (10)$$

where the risk reduction,  $e_c$ , is in the interval  $[e_1, e_2]$ . Moreover, letting  $w_{e_c} = w - P(e_c, w, \pi)$  we obtain:

$$\frac{\partial P(e_c, w, \pi)}{\partial e} = \frac{u_A(w_{e_c}) - u_D(w_{e_c})}{(\pi + e_c) u'_A(w_{e_c}) + (1 - \pi - e_c) u'_D(w - P(e_c, w, \pi))} > 0.$$

This can be re-expressed as

$$\frac{\partial P(e_c, w, \pi)}{\partial e} = \frac{P(e_2, w, \pi) - P(e_1, w, \pi)}{e_2 - e_1} = \frac{\partial P(0, w_{e_c}, \pi + e_c)}{\partial e} \equiv VSL_{e_c}, \quad (11)$$

where  $VSL_{e_c}$  denotes the value per statistical life at wealth  $w_{e_c}$  and baseline survival  $\pi + e_c$ . It follows from (11) that  $VSL_{e_c}$  is also a valid measure of VSL. We describe  $VSL_{e_c}$  as a "scope-revealing VSL" (SR-VSL). In the CV literature, (8) and (9) (or their average) are most commonly reported. To our knowledge, (11) has not been previously described. Equation (11) is equivalent to the incremental willingness to pay derived from a discrete choice setting where an individual is indifferent between paying for two alternative risk reductions: a risk reduction  $e_1$  at price  $P(e_1, w, \pi)$  and a risk reduction  $e_2$  at price  $P(e_2, w, \pi)$ .

Theoretically, the differences between the three measures of VSL are minimal for small risk reductions (in the limit as  $e_2 \rightarrow 0$  they are equal). Empirically, a large majority of contingent valuation studies addressing mortality risk reductions suffer from inadequate scope sensitivity (WTP is increasing in the risk reduction but less than proportionate). When a survey yields estimates of WTP that are not proportional to the risk reduction  $e$ , there are several estimates of VSL that can be reported, including  $VSL_{e_1}$ ,  $VSL_{e_2}$ , and their average. We propose a supplemental measure,  $VSL_{e_c}$ . When WTP is less than proportionate to the risk reduction, by equations (8), (9) and (11),  $VSL_{e_1} > VSL_{e_2} > VSL_{e_c}$ .  $VSL_{e_1}$  and  $VSL_{e_2}$  are direct estimates of VSL, in that each is derived using stated WTP for that exact risk change; their average is a plausible direct estimate of the VSL from the study. The SR-VSL,  $VSL_{e_c}$ , tends to zero as scope sensitivity tends to zero. Hence the difference between  $VSL_{e_c}$  and the average of  $VSL_{e_1}$  and  $VSL_{e_2}$  is a measure of how much the results differ from perfect scope sensitivity.

A less-than-proportionate relationship between WTP and risk reduction may arise from the curvature of the utility function for wealth (i.e., the income effect), from respondents' limited understanding of the good being valued (Hammit and Graham, 1999), or from

other factors.<sup>12</sup> In designing questions to elicit the marginal valuation of risk reduction, there is a trade-off between distortions due the curvature of the utility function and to respondents' comprehension of the risk reduction. WTP for small risk reductions yield a more accurate measure of the marginal rate of substitution between wealth and risk in principal, but as risks become smaller the cognitive difficulty for respondents may increase. For small risk reductions, the curvature problem is of second order while respondents' understanding remains a first order issue.

### 3 Survey design and descriptive statistics

This section describes the survey instrument, administration, and descriptive statistics.

#### 3.1 Structure of the questionnaire and survey administration

The survey was designed to elicit WTP for a reduction in the probability of death, as in the theoretical model. The source of mortality risk is pesticide residues on food, which can be reduced by purchasing an alternative food produced using safer but more expensive pesticides. The survey instrument was virtually identical, save for language, to that used by Hammitt & Haninger (2010) in the US. The survey was administered to a random sample of the CSA Online Panel (CSA is a French survey company). Panel members were recruited by random sampling using email and closely match the French adult population on age, gender, socio-economic and geographical variables. Respondents were compensated with points that can be used to purchase items online. Data were gathered in two waves between July and August 2012, with 1000 surveys completed. Our final sample of adult respondents includes 192 from single-person households, 296 from households that include at least one other adult and no child, 130 from households that include no other adult and at least one child, and 382 from households that include at least one other adult and one child. The survey instrument was divided into several sections: introduction, verification of household composition, practice questions, valuation, and general follow-up questions.

In the introduction, respondents were informed about the objective of the survey, the approximate time it would take to complete, and the number of points awarded for participation. Upon agreement to participate, they were asked about the number of children younger than 18 years and the number of other adults living in their household.

Following a warm-up question about the respondent's concern for air pollution, risks on the job, pesticides in food, automobile accidents, and airplane crashes, respondents were presented with two practice valuation questions with feedback. In the first, respondents were asked to choose between two types of apples: one type had both a lower probability of causing illness and a lower price than the other. Respondents who chose the dominant alternative were told that the type they had selected was both safer and less expensive and

---

<sup>12</sup>When non-satiation is violated, scope sensitivity may not be required (Banerjee et al. 2005). Also, results may not exhibit sufficient sensitivity to scope if WTP is elicited using open-ended valuation questions and respondents round their responses (Whynes et al. 2005).

that this was the logical choice. Respondents who chose the dominated alternative were told that the type they had selected was both less safe and more expensive than the other and asked if they wished to choose again. In the second practice question (about grapes), neither alternative was dominant. Respondents were told the type they had chosen was safer and more expensive, or less safe and less expensive, as appropriate and asked to confirm that was the choice they preferred.

Following the practice questions, respondents were asked about the age, sex, and current health of each of the possible targets of a risk reduction: the respondent him or herself, a child and another adult living in the household (when a household included multiple children or other adults, the respondent was asked to choose the one whose birthday came next). Current health was elicited using a visual analog scale on which 100 corresponds to full health and 0 to a state as bad as death (which we divided by 100 to yield a value between 0 and 1), and using the EQ-5D health state classification system (Euro-Qol Group 1990). The EQ-5D is a utility instrument used to estimate the health-related quality of life (HRQL) associated with an individual's current health or a hypothetical health state. It classifies health states using five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression), each of which can take any of three levels (no, moderate, or severe problems). HRQL is an index normalized to a value of one for full health and zero for health states an individual judges indifferent to dead (negative values are permitted). HRQL can be estimated from the EQ-5D health-state description using a scoring rule; as no scoring rule is available for France, we use the scoring rule developed by Shaw et al. (2005) using preferences for health states elicited from a large representative sample of the US population.

The valuation section was divided into subsections concerning each of the targets about which a respondent was questioned: the respondent, a child and another adult (if present). The order of targets was randomized. For each target, the respondent was asked to suppose that individual had a specified chance of developing a fatal chronic disease caused by exposure to pesticides in food. The risk could be reduced by purchasing an otherwise identical food produced through a hypothetical "Pesticide Safety System" that used alternative pesticides which are safer to humans (i.e., the alternative is not organically grown food). The description emphasized that the food types did not differ in taste, appearance, or nutritional value, to guard against the chance respondents would infer there were other benefits of purchasing the more expensive food. Respondents were asked to assume that only the targeted individual would eat the food in question so that responses can be interpreted as reflecting WTP to reduce risk to a defined individual. As food is often shared within a household, this assumption is not always realistic and failure of respondents to accept it could bias our estimates. However, other studies have used a similar scenario and found apparently reasonable differences in WTP to reduce risk to different targets (e.g., Hammitt and Haninger (2007, 2010)). Following the questions about individual targets, respondents were asked about WTP to reduce risk to all household members simultaneously (i.e., assuming everyone would eat the same food). Depending

on household composition, respondents answered between two and four valuation questions (risks to self, household, a child and another adult if present). Respondents from a single-person household were asked valuation questions about two risks to themselves.

The disease, baseline risk, risk reduction, and latency were randomly varied between valuation questions. The disease was described by the organ affected (bladder, brain, liver or white blood cells) and whether it was cancer or not. Half the respondents were also provided with a few paragraphs describing the disease symptoms (see appendix).<sup>13</sup> The latency period between exposure and when symptoms would first appear was described as 1, 10, or 20 years. After the first symptoms appeared, the morbidity would persist for two years, followed by death. For each scenario, the respondent was asked to judge the target's health while ill using both the visual analog scale and the EQ-5D. By asking the respondent to evaluate health conditional on having the disease immediately prior to the valuation question we attempted to focus his or her attention on the characteristics of the disease risk to be reduced. The baseline risk (3 or 4 per 10,000 per year) and risk reduction (1 or 2 per 10,000 per year) were illustrated using a visual aid (Corso et al. 2001) in which areas of the computer screen proportional to these probabilities and to the complementary probability of no illness were distinctively colored. A sample valuation question is shown in Figure 1.

In each valuation question, the initial risk, risk reduction using the alternative food type, and additional annual cost of the alternative food type were specified and the respondent asked to choose which food type she or he would select. Values were elicited using a standard double-bounded binary-choice format (Hanemann et al. 1991). The initial bid (the incremental cost of the safer food type) varied between € 10 and € 6,000 per year; the follow-up bid was twice the initial bid for respondents who indicated they would choose the safer food in the initial question and half the initial bid for other respondents.

Respondents were asked a series of questions about their degree of confidence in their responses, whether they thought the risks faced by targets were consistent with those stated in the survey, and whether they interpreted the scenario in which risk to all household members was reduced as implying that all household members would (or would not) develop the disease. The survey ended with additional questions concerning food preferences, purchase patterns and demographics (income and education).

### 3.2 Data

Tables 2 and 3 provide descriptive statistics on the sample of households that completed the survey. Sample means and standard deviations are reported for the entire sample and for each subsample of respondents who answered questions about risk to a child or to another adult living in the respondent's household. The average age of a respondent is 43 years; half of respondents are female and half have at least a bachelors degree. The children are 9 years old on average, and the other adults average 42 years old. The net

---

<sup>13</sup>Six physicians from a range of specialties reviewed and helped refine the descriptions.

(A) Introduction to a valuation question

Next, we would like to ask how much it would be worth to you to reduce your chance of developing [disease name] by being exposed to pesticides in food. Suppose it would be possible to reduce this chance if the food producer were to follow a stringent safety program that is established and monitored by the United States Government. This program is called "Pesticide Safety System" and includes the use of pesticides that are safer to humans than conventional pesticides. While the food produced by the Pesticide Safety System is safer to humans than conventional food, the Pesticide Safety System is not an organic farming practice, nor does it affect the environment any differently than conventional farming.

**[Next Screen]**

Because it requires more expensive pesticides, a producer that follows the Pesticide Safety System would need to charge a higher price. Some producers will adopt the Pesticide Safety System, and others will not. You would be able to tell easily, from a label or sticker, whether the food you are thinking of buying was produced by conventional farming or by the Pesticide Safety System. Whenever you go to the store, you would be able to choose whether to buy the same food you buy now, or to buy food produced by the Pesticide Safety System at a somewhat higher price. Food produced by the Pesticide Safety System does not differ in taste, appearance, or nutritional value.

(B) Valuation question without symptom description

Whether you eat the Pesticide Safety System food or the conventional food, if you develop [disease name], you would not be diagnosed or begin to experience symptoms for [20 years].

After that time, you could expect to be sick for about two years and then die.

The following table summarizes the differences between the Pesticide Safety System food and the conventional food:

Type of Food	Your Chance of Developing [Disease Name] by Being Exposed to Pesticides in Food	Annual Cost
Pesticide Safety System	[2 in 10,000] per year	[X] more per year than conventional food
Conventional	[4 in 10,000] per year	conventional cost

If Pesticide Safety System food cost [X] more per year than conventional food, which type of food would you purchase?

- Pesticide Safety System
- Conventional

*Notes:* Text and numbers in brackets are randomly assigned. Next screen in brackets refers to displaying the text situated below in a separate screen.

Figure 1: Typical valuation question.

monthly income of the average household is 2910 € (in 2012 €).<sup>14</sup>

The share of respondents that chose the dominant alternative in the practice question is 80%. About 66% of the respondents were somewhat or very confident in their answers to the valuation questions and approximately 67% thought that the actual risks faced by the targets were about the same as the risks presented in the scenarios. After being presented with the household scenario, only 10% of respondents thought that if one member of the household developed the disease, all the other members would also. Most of the respondents (80%) reported that they purchase the food for the household all or most of the time. Purchase of bio-labeled (i.e., "organic") food is less frequent with about 60% purchasing it occasionally, rarely, or never. Respondents typically prepare their own meals and are somewhat concerned about pesticides in foods. Finally, respondents spent a median of 18 minutes completing the survey.

Current health, as measured with a visual analogue scale and EQ-5D, is similar for self and other adults and higher for children. HRQL elicited using EQ-5D is consistent with the measurement elicited with the visual analogue measure. Anticipated health conditional on developing one of the stated illnesses is about 40% lower using both measures. The fraction of answers reporting a higher HRQL in the sick state than in the healthy state is 16% for VAS and 2.5% for EQ-5D. Respondents who were presented with the supplemental disease descriptions judged the disease as more severe, with smaller HRQL if ill.

## 4 Empirical implementation

We analyze respondents' WTP using both standard and LCA models, as described in the following subsections. In both cases, the dependent variable is the natural logarithm of WTP, which is interval-censored because it is elicited using double-bounded dichotomous-choice questions (Hanemann et al. 1994).

### 4.1 Standard analysis

Assume that each of  $N$  respondents answer valuation questions concerning different risks or targets on  $T$  choice occasions. Define the WTP,  $P(e, w, \pi)$ , of respondent  $i$  on choice occasion  $t$  for a risk reduction  $e$  as:

$$\log(P_{it}(e, w_i, \pi)) = \beta_1 \log(RR_{it}) + \beta_2 \log(BLR_{it}) + \beta_3 \log(INC_i) + z_{it}\beta_4 + \xi + \epsilon_{it}, \quad (12)$$

where  $RR_{it}$ ,  $BLR_{it}$  and  $INC_i$  correspond to risk reduction, baseline risk, and income, respectively;  $z_{it}$  contains other individual characteristics including target dummies (child, other adult, and household); and  $\xi$  is a constant. All remaining heterogeneity is captured

---

<sup>14</sup>About 16 percent of households failed to report their income. A probit analysis suggests that household composition is the only significant predictor of income being missing (single-person households were less likely to report income). Missing values were imputed as the average conditional on the type of household. Estimates of the relationship between household income and WTP (described below) are not significantly affected by imputing the overall mean or deleting observations with missing income.

Table 2: Summary statistics, Panel A

	Pooled	Self	Child	O. Adult
Age of person at risk	37.22 (17.47)	42.91 (13.14)	9.07 (5.09)	41.64 (13.89)
Female	0.51 (0.50)	0.52 (0.50)	0.47 (0.50)	0.51 (0.50)
Education	0.78 (0.41)	0.78 (0.41)	0.80 (0.40)	0.78 (0.42)
Household income (€/month)	2980 (1708)	2909 (1716)	2839 (1570)	3299 (1752)
Did not fail training	0.82 (0.38)	0.82 (0.38)	0.81 (0.38)	0.81 (0.38)
Respondent somewhat confident	0.66 (0.47)	0.66 (0.48)	0.64 (0.48)	0.68 (0.47)
Hypothetical and actual risk are similar	0.65 (0.48)	0.67 (0.47)	0.62 (0.49)	0.66 (0.47)
If a member gets sick, all get sick?	0.10 (0.32)	0.10 (0.32)	0.14 (0.34)	0.11 (0.31)
Frequency purchase food for household	0.86 (0.35)	0.86 (0.34)	0.89 (0.31)	0.81 (0.39)
Frequency purchase BIO	0.63 (0.48)	0.64 (0.48)	0.62 (0.49)	0.61 (0.49)
Frequency preparing meals	0.78 (0.41)	0.78 (0.41)	0.82 (0.38)	0.76 (0.43)
Pesticides in food	0.60 (0.49)	0.60 (0.49)	0.59 (0.49)	0.61 (0.49)
Log-time spent filling survey (seconds)	7.29 (1.32)	7.25 (1.29)	7.34 (1.38)	7.34 (1.36)
N	3,190	1,000	512	678

*Notes:* Female is a dummy variable taking the value 1 when female, 0 otherwise. Education is a categorical variable taking values of 0 when no high school degree, 1 when high school degree, college degree or higher. Wage corresponds to household net income in 2012 €. Did not fail training takes value 1 when the respondent did not fail during the training, 0 otherwise. Confidence in answering valuation question is a self reported variable taking values 1 (very confident), 2 (somewhat confident) 3 (not at all confident); the value reported in the table correspond to share of respondents answering value 2. Hypothetical and actual risk similarity question is a self reported value taking a value 1 (the chance is higher), 2 (about the same) and 3 (the chance is lower); the variable used in the paper corresponds to the share of respondents answering value 2. If a member gets sick, all get sick? corresponds to the answer of a follow-up question after the household risk scenario, it takes value 1 when the answer is no and 0 otherwise. Frequency of purchase of food for household question is a self reported value taking a value 1 (very often), 2 (nearly all the time), 3 (half of the time) 4 (Occasionally) and 5 (Rarely or never); the variable used in the paper corresponds to the share of respondents answering values 1 and 2. Frequency of purchase of Bio labeled food is a self reported value taking the same values as the purchase for household food; the variable used in the paper corresponds to the share of respondents answering values 4 and 5. Frequency of preparing own meals is a self reported value taking the same values as the purchase for household food; the variable used in the paper corresponds to the share of respondents answering values 1 and 2. Pesticides in food is a self reported variable taking values from 1 (low) to 5 (high); the variable used in the paper corresponds to the share of respondents answering values 4 and 5.

Table 3: Summary statistics, Panel B

Adult health	Average	Description		<i>P</i> -value
		Yes	No	
Current health (EQ-5D)	0.87	0.87	0.87	0.463
Health in illness (EQ-5D)	0.50	0.45	0.55	0.01<
Loss in health (EQ-5D)	0.37	0.42	0.32	0.01<
Current health (VAS)	0.78	0.78	0.78	0.563
Health in illness (VAS)	0.49	0.47	0.51	0.036
Loss in health (VAS)	0.29	0.30	0.27	0.108
Child health	Average	Description		<i>P</i> -value
Current health (EQ-5D)	0.95	0.96	0.95	0.429
Health in illness (EQ-5D)	0.48	0.43	0.52	0.01<
Loss in health (EQ-5D)	0.48	0.52	0.43	0.01<
Current health (VAS)	0.91	0.91	0.91	0.731
Health in illness (VAS)	0.49	0.45	0.52	0.01<
Loss in health (VAS)	0.42	0.45	0.38	0.01<
Other adult health	Average	Description		<i>P</i> -value
Current health (EQ-5D)	0.88	0.88	0.87	0.239
Health in illness (EQ-5D)	0.51	0.48	0.54	0.01<
Loss in health (EQ-5D)	0.37	0.40	0.33	0.01<
Current health (VAS)	0.79	0.80	0.79	0.570
Health in illness (VAS)	0.53	0.51	0.55	0.043
Loss in health (VAS)	0.26	0.29	0.23	0.033

*Notes:* Current health state is a self-reported measure of current health ranging from 0 to 100, respectively. Disease health state is a self-reported measure of health when sick with the disease described in the survey ranging from 0 to 100. EQ-5D score for illness is computed using standard weights. Loss in health when (not) described corresponds to the average loss in health conditional on (not) describing the symptoms. *P*-value tests each variables' difference between description of the disease and no description of the disease.

by the idiosyncratic shocks  $\epsilon_{it}$ , which are assumed to be independently drawn from a normal distribution.

The coefficients of the models are estimated using maximum likelihood estimation (Alberini 1995) and the standard errors are calculated using a Wald test (Train 2009). We allow for correlation between errors within a respondent across choice occasions, but assume independence between respondents.

From the regression (12), we can predict the logarithm of respondent  $i$ 's WTP on choice occasion  $t$ ,  $\log(\hat{P})$ , by setting the right-hand-side variables to their respondent- and choice-occasion-appropriate values and  $\epsilon_{it} = 0$ . To estimate VSL, we must retransform  $\log(\hat{P})$  to euros then divide by the risk reduction. Many CV studies estimate  $P$  as  $\exp(\log(\hat{P}))$ , which is the predicted median over the error term. Alternatively, one may estimate the predicted mean over the error term by  $\exp(\log(\hat{P}) + \frac{\hat{\sigma}^2}{2})$ , where  $\hat{\sigma}^2$  is the estimated variance of  $\epsilon_{it}$ . These alternatives yield two estimates of VSL,

$$VSL_{mean} = \frac{1}{e} \exp\left(\log(\hat{P}) + \frac{\hat{\sigma}^2}{2}\right); \quad VSL_{med} = \frac{1}{e} \exp\left(\log(\hat{P})\right); \quad \frac{VSL_{mean}}{VSL_{med}} = \exp\left(\frac{\hat{\sigma}^2}{2}\right).$$

The two values differ by a factor of  $\exp(\frac{\sigma^2}{2})$ , which can be an order of magnitude or more when there is substantial variation around the estimated regression equation. If the error term captures only independent and identically distributed random error, this specification is appropriate. However, to the extent that the regression fails to account for systematic variation across respondents or choice occasions, it overstates the error in predicting the logarithm of WTP and biases the predicted mean WTP upward. For example, if some of the heterogeneity arises from unobserved respondent characteristics, including perhaps differences in comprehension of the survey questions or the effort they devote to answering the questions, predicted mean WTP is biased upward.

Duan (1983) notes that if the error in eqn. (12) is not distributed normally, an estimate of the mean of  $P$  that assumes normality, and hence of  $VSL_{mean}$ , will not be consistent. He proposed a smearing estimate of the expected response on the untransformed scale using the empirical distribution of the error. With interval-censored data, the empirical distribution of the error is poorly estimated; hence we do not attempt to apply the smearing estimator in this context.

## 4.2 Latent Class Analysis

Unobserved individual heterogeneity abounds in contingent valuation studies: individuals differ in their cognitive resources and may differ in their attention to different aspects of the survey questions (Cameron & DeShazo 2013). LCA is a valuable method to assess unobserved heterogeneity (Train 2009). In a recent paper, Hess et al. (2011) suggest latent class models are able to uncover richer patterns of heterogeneity than continuously mixed models. We assume that the underlying coefficients in the WTP function follow a discrete

distribution, which LCA estimates non-parametrically. Hence, LCA enables us to group together respondents that have similar preference structures.

Consider a respondent  $i$  who belongs to class  $s$  where  $s = 1, \dots, C$  and  $C$  is the number of classes. Respondent's  $i$  WTP on choice occasion  $t$  for a risk reduction  $e$  is:

$$\log(P_{it}(e, w_i, \pi)) = \beta_{1s} \log(RR_{it}) + \beta_{2s} \log(BLR_{it}) + \beta_{3s} \log(INC_i) + z_{it} \beta_{4s} + \xi_s + \epsilon_{ist}, \quad (13)$$

where  $\xi_s$  corresponds to a constant unobservable class  $s$  fixed effect and  $\epsilon_{ist}$  rationalizes all remaining choice-to-choice individual variation. Notice that all coefficients in equation (13) are class-specific.

Let  $b_{it0}$  represent the initial log-bid for individual  $i$  at choice  $t$ ,  $b_{itU}$  the follow-up log-bid if the individual opts in favour of the risk reduction and  $b_{itL}$  otherwise. Moreover, let  $x_{1it} = \{\log(RR_{it}), \log(BLR_{it}), \log(INC_{it}), z_{it}\}$  and  $x_{it} = \{x_{1it}, x_{2it}\}$  represent a matrix of size  $N \times (K1 + K2)$  of individual characteristics. The matrix is divided between characteristics that explain WTP,  $x_{1it}$ , and characteristics that explain the probabilities of membership in each class,  $x_{2it}$ , which may or may not overlap.

We assume  $\epsilon_{ist}$  is distributed normally. Hence, the conditional probability that WTP of individual  $i$  belongs to a particular interval is given by:

$$Q_{it}(\theta_s, x_{1it}, y_{it}) = \begin{cases} \Phi\left(\frac{b_{itL} - x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it} = 0 \\ \Phi\left(\frac{b_{it0} - x_{1it}\beta_s}{\sigma_s}\right) - \Phi\left(\frac{b_{itL} - x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it} = 1 \\ \Phi\left(\frac{b_{itU} - x_{1it}\beta_s}{\sigma_s}\right) - \Phi\left(\frac{b_{it0} - x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it} = 2 \\ 1 - \Phi\left(\frac{b_{itU} - x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it} = 3 \end{cases}, \quad (14)$$

where  $\Phi$  is the standard normal cumulative distribution function and  $\theta_s = (\beta_s, \sigma_s)$  are the mean and residual error parameters of the normal distribution for the class  $s$ . The choice indicator  $y_{it}$  represents responses to the initial and follow-up binary-choice questions of "No-No", "No-Yes", "Yes-No", and "Yes-Yes", respectively. Since  $\theta_s$  is unknown, the sequence of observed choices must be evaluated over all possible values. We assume that the density of the parameters is described by a discrete distribution. It follows that the log-likelihood function is:

$$LL(\Theta) = \sum_{i=1}^N \log \left( \sum_{s=1}^C \pi_{is}(x_{2it}, \alpha_s) \prod_{t=1}^T Q_{it}(\theta_s, x_{1it}, y_{it}) \right), \quad (15)$$

where  $\Theta = (\theta_1, \dots, \theta_C; \alpha_1, \dots, \alpha_C)$  comprises all model coefficients,  $\pi_{is}(x_{2it}, \alpha_s)$  correspond to the prior probabilities of individual  $i$  belonging to class  $s$ , and  $\alpha_s$  corresponds to the association of demographics  $x_{2it}$  with class membership  $s$ . To better understand, the log-likelihood can be re-expressed as follows:

$$LL(\Theta) = \sum_{i=1}^N \log(L_i^s),$$

where,

$$L_i^s = \pi_{is}(x_{2it}, \alpha_s) \prod_{t=1}^T Q_{it}(\theta_s, x_{1it}, y_{it}).$$

The main identifying assumption is that respondents' unobserved shocks are independent between respondents (Train 2009). In principle this function can be maximized through full information maximum likelihood, but in general it is easier to do with an Expectation Maximization algorithm (Dempster et al. 1977). The problem, which is solved with EM, is that class membership is missing and has to be estimated. Notice that if we knew the number of classes, and to which class each agent belongs, we would have to estimate C conventional likelihoods.<sup>15</sup>

## 5 Results

### 5.1 Standard analysis

In Table 4 we report estimates of equation (12). Model (1) is estimated using the responses to all valuation questions and examines the effects of risk reduction, baseline risk, and income on WTP. Of these, only the coefficient on risk reduction is significantly different from zero. It is also significantly different from 1, violating our *e*-test. Individuals are willing to pay 1.35 ( $= \exp(0.436 \log(2))$ ) times more for a risk reduction of 2 in 10,000 than a risk reduction of 1 in 10,000. The estimated coefficient on baseline risk is not significantly different from zero, consistent with our  $\pi$ -test. The coefficient on income is not significantly positive, but because it is not significantly negative it satisfies our *w*-test. Respondents living in a household with more than one member are willing to pay 1.5 ( $= \exp(0.406)$ ) times more for a risk reduction to all members of the household (including themselves) than to themselves alone. Respondents living alone report consistent willingness to pay between the two risk reduction questions. Finally, respondents are willing to pay on average 2.6 ( $= \exp(0.974)$ ) times more to avoid a risk to their child than to themselves and 2.0 ( $= \exp(0.686)$ ) times more to avoid a risk to another adult in their household.

Model (2) in Table 4 is estimated using only answers to the first valuation question presented to each respondent, plus the last question which concerned reducing risk to all household members jointly (or to the respondent alone in a single-person household). The coefficient on risk reduction is significantly different from zero and not from one, satisfying our *e*-test. Respondents are willing to pay 1.6 ( $= \exp(0.69 \log(2))$ ) times more for a risk reduction of 2 in 10,000 than for a risk reduction of 1 in 10,000. The larger coefficient on risk reduction as compared with model (1) might be attributed to respondents' fatigue or lack of motivation in responding to second and third valuation questions.<sup>16</sup>

<sup>15</sup>A detailed explanation of how the EM-algorithm is implemented can be found in the appendix.

<sup>16</sup>Of the 11.5% of respondents who were presented with the same initial bids and risk reductions for two valuation questions (concerning different diseases or targets), almost 70% gave the same response to the initial bid. These responses are potentially explained as examples of arbitrary coherence (Ariely et

As with model (1), both the  $\pi$ -test and the  $w$ -test are satisfied. A respondent living in a household with more than one member is willing to pay 1.9 ( $= \exp(0.652)$ ) times more to reduce a risk to all members than to reduce a risk to him or herself. Respondents living in a single-person household answered the two WTP questions consistently (the difference between the first and second questions is not significantly different from zero). Finally, respondents are willing to pay 2.6 ( $= \exp(0.96)$ ) and 2.5 ( $= \exp(0.918)$ ) times more to reduce risk to a child and to another adult than to themselves, respectively. For comparison, Hammitt and Haninger (2010) found smaller differences in WTP to reduce risk to a child or to another adult, compared with risk to self. They estimated ratios of 1.8 and 1.15, respectively; their estimate for another adult is not statistically significantly different from one.

Models (3) and (4) include the same variables and observations as models (1) and (2), respectively. In addition, they include disease characteristics. Coefficients relevant to our validity criteria are not affected by adding these variables. None of the coefficients of the disease characteristic variables are significantly different from zero. This implies that WTP is not sensitive to whether the disease is cancer or not, whether the latency period is 1, 10, or 20 years, and whether it affects the brain, liver, white blood cells, or bladder. Similarly, Hammitt and Haninger (2010) found no significant effects of disease type, affected organ, or latency. Note that standard theory does not require that WTP decrease with latency (Hammitt and Liu 2004).

Table 5 reports the median VSL for models (1) and (2), for each target individual and disease. These values are predicted medians over the error term; they are for the sample-mean respondent. The interpretation of expressions (8), (9), and (11) as measures of the marginal rate of substitution of wealth for risk depends on two assumptions: first, that the risk change is close to zero; second, that there is perfect scope sensitivity (WTP is proportional to risk reduction). In the case of (11), the risk reduction is small enough as it is the difference  $e_2 - e_1$ . Under perfect scope sensitivity, SR-VSL computed from equation (11) is equal to the values from expressions (8) and (9); if WTP is less than proportional to the risk reduction, SR-VSL is smaller than the conventional estimates, revealing the lack of scope sensitivity in elicited WTP.

It is clear from table 5 that the conventional VSL calculated using expression (8) or (9) (but for the mean risk reduction, 1.5/10,000) and SR-VSL do not coincide. For the respondent him or herself, VSL is approximately 4 to 5 million €, while the SR-VSL is approximately 2 to 3 million €. For a child, the VSL is about 10 to 13 million € while SR-VSL is around 6 to 7 million €.

Table 5 also reports VSL estimates for the respondent when different organs are affected, latency and cancer status are changed, and whether a symptom description is given or not. We do not find a statistically significant difference between VSL for the different disease scenarios. Also, we cannot reject the hypothesis that SR-VSL from model 1 is

---

al. 2003), but the log risk reduction coefficient is unchanged when we exclude these respondents from the sample.

Table 4: Willingness to pay results: Standard analysis

	Model (1)	Model (2)	Model (3)	Model (4)
Log-risk reduction	0.436** (0.220)	0.690** (0.285)	0.434** (0.219)	0.695** (0.284)
Log-baseline risk	0.185 (0.515)	0.195 (0.670)	0.181 (0.514)	0.198 (0.669)
Log-income	-0.0006 (0.238)	0.0362 (0.242)	0.0006 (0.239)	0.0250 (0.243)
Child is at risk	0.974*** (0.167)	0.960** (0.380)	0.669*** (0.141)	0.918*** (0.288)
Adult is at risk	0.686*** (0.139)	0.918*** (0.287)	0.962*** (0.167)	0.970** (0.380)
Household is at risk	0.406*** (0.123)	0.652*** (0.202)	0.379*** (0.128)	0.649*** (0.207)
Self is at risk (2d question)	-0.197 (0.284)	0.0625 (0.273)	-0.228 (0.285)	0.0678 (0.276)
Brain			0.0555 (0.149)	0.121 (0.234)
Liver			-0.127 (0.148)	-0.131 (0.239)
White Blood Cells			0.0628 (0.148)	0.314 (0.241)
Latency is 10 years			0.0324 (0.177)	-0.0691 (0.234)
Latency is 20 years			0.0252 (0.183)	-0.106 (0.234)
Cancer			0.195 (0.143)	0.166 (0.192)
Symptom description			-0.0577 (0.231)	-0.133 (0.240)
Constant	11.95** (5.051)	13.73** (6.186)	11.84** (5.047)	13.81** (6.177)
Sigma	3.396*** (0.147)	3.495*** (0.158)	3.395*** (0.145)	3.488*** (0.158)
Observations	3,190	2,000	3,190	2,000

*Notes:* Dependent variable is an interval censored log-WTP, measured using a double-bounded elicitation method. Follow up bids are double or halved, if the respondents agree, or disagree, to pay the initial bid. Respondents answers to WTP for each risk reduction in the study are pooled. Respondents idiosyncratic shocks are allowed to be correlated between questions. The log-risk reduction variable takes the value of  $\log(1/10,000)$  if the respondents are faced with 1/10,000 with a risk reduction and takes the value of  $\log(2/10,000)$  if the respondents are faced with a 2/10,000 risk reduction. The log baseline risk variable takes a value of  $\log(4/10,000)$  if the baseline risk is 4/10,000 and  $\log(3/10,000)$  otherwise. The order of the person to which the risk reduction was addressed is random. Model (2) and (4) report results over the subset of answers to the first valuation question, and questions on household level risk including second questions to single person households. The household WTP question is always asked last. Standard errors clustered at the respondent level are in parenthesis.

\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

equal to SR-VSL from model 2.

Table 5: Median value per statistical life

	Model (1)		Model (2)	
	VSL	SR-VSL	VSL	SR-VSL
VSL for self	4.92 (0.62)	2.24 (0.30)	3.77 (0.70)	2.74 (0.54)
VSL for child	13.02 (2.36)	5.93 (3.83)	9.86 (3.26)	7.16 (9.95)
VSL for another adult	12.57 (2.19)	4.44 (1.68)	9.45 (2.03)	6.86 (4.29)
VSL for per HH member	2.53 (1.10)	1.15 (0.90)	2.48 (1.11)	1.80 (1.50)
VSL for self (2de question)	4.04 (1.20)	1.84 (0.90)	4.02 (1.21)	2.92 (1.39)
<b>VSL for self when:</b>				
Brain is affected	5.28 (0.86)	2.39 (0.52)	3.93 (0.97)	2.87 (0.94)
Bladder is affected	5.00 (0.80)	2.26 (0.45)	3.48 (0.83)	2.55 (0.70)
Liver is affected	4.40 (0.65)	1.99 (0.31)	3.05 (0.72)	2.23 (0.52)
White blood cells are affected	5.32 (0.80)	2.41 (0.47)	4.76 (1.08)	3.48 (1.20)
Disease latency is 1 year	4.89 (0.81)	2.21 (0.46)	4.00 (0.94)	2.93 (0.89)
Disease latency is 10 years	5.05 (0.81)	2.29 (0.46)	3.73 (0.85)	2.73 (0.74)
Disease latency is 20 years	5.01 (0.83)	2.27 (0.48)	3.60 (0.83)	2.63 (0.70)
Disease is not cancer	4.52 (0.64)	2.05 (0.31)	3.47 (0.71)	2.54 (0.53)
Disease is cancer	5.49 (0.82)	2.49 (0.49)	4.09 (0.89)	3.00 (0.82)
Symptom description provided	4.84 (0.86)	2.19 (0.51)	3.53 (0.80)	2.58 (0.65)
Symptom description not provided	5.13 (0.85)	2.32 (0.51)	4.03 (0.88)	2.95 (0.81)

*Notes:* Values are in millions of €. log-WTP is calculated using the specification from each model. Median VSL is estimated for the mean individual in the following way: first, we take the exponential log-WTP for the mean individual at the average risk reduction 1.5/10000 risk reduction; second, we divide by the average risk reduction (1.5/10,000). Median SR-VSL is estimated for the mean individual in the following way: first, we take the exponential log-WTP for the mean individual when presented with a 2/10000 risk reduction and a 1/10000, separately; second, take the difference between the exponential log-WTP when risk reduction is equal to 2/10000 and the exponential log-WTP when risk reduction is equal to 1/10000; third, divide the result from the second step by the difference between high and low risk reduction. Standard errors are in parenthesis (delta method).

## 5.2 Latent class analysis

Latent class analysis (LCA) provides a method to better understand the heterogeneity among respondents. LCA endogenously creates groups of homogeneous classes of respon-

dents, where each has a positive, respondent-specific probability of membership in each class. We use the term "class X members" to describe the weighted contribution of inputs from all respondents with weights equal to the posterior probabilities of membership in class X.

We perform a LCA for  $C = 2, \dots, 5$  classes. The preferred model, given the Bayesian Information Criterion, is the model with  $C = 3$  classes. The first, second and third classes have average posterior membership probabilities of 29%, 42% and 29%, respectively. In Table 6 we report estimated coefficients for the 3-class model as well as posterior probability estimates. Class membership is estimated using responses to all valuation questions but we estimate WTP for only the first target and the household (as for model (4) in the standard analysis).

First, consider the posterior coefficients (Table 6, last column). These can be compared directly with the estimated coefficients of the standard model (Table 4, model (4)), which has the same covariates (Train 2009). The posterior coefficients are constructed as the sum of the coefficients estimated for each class, weighted by the class's posterior probability. The posterior coefficient on log-risk reduction is statistically different from zero and not from one, which satisfies our e-test. The relationship between baseline risk and WTP is negative but not significant, which is consistent with our  $\pi$ -test. Elasticity of WTP with respect to income is positive but not significantly different from zero, satisfying our  $w$ -test. Estimates of the remaining coefficients are not significantly different from the estimates from model (4) in Table 4. Noisy estimates of posterior coefficients are to be expected (Train 2009).

Next consider the coefficients for the class-specific regression. Estimates for class 2 satisfy our primary validity tests; in contrast, estimates for classes 1 and 3 violate one or more of these tests. For class 2, the estimated coefficient on log risk reduction, 0.74, is significantly different from zero and not from one, which is consistent with proportionality and satisfies our e-test. The coefficients on baseline risk and income are both positive but not significantly different from zero, satisfying our  $\pi$ -test and  $w$ -test, respectively. Median (mean) WTP (median = 31 €, mean = 261 €) does not exceed the average Class 2 income, (2878 €).

Class 2 estimates show that the value of reducing risk to a child is 2.7 times ( $\exp(1.02)$ ), and to another adult is about 2.1 times ( $\exp(0.754)$ ), larger than the value of reducing risk to the respondent him or herself. However, WTP to reduce a risk to the entire household is not statistically significantly different than WTP to reduce risk to the respondent alone, violating our test that WTP to reduce risk to the household exceeds WTP to reduce risk to a single member. For respondents in single-person households, there is no significant difference in WTP between the first and second valuation questions, consistent with our criterion.

Finally, WTP is sensitive to information about the disease. WTP is significantly smaller when the symptom description is provided and when the latency period is 20 years, but not when the latency is 10 years (compared with one year). The implied

discount rate for the 20 year latency is 2.7% per year. Other disease characteristics are not statistically significant, except WTP is smaller when the affected organ is the liver.

For class 1, the coefficient on risk reduction is not significantly different from zero but is different from one, violating our  $e$ -test. The coefficient on baseline risk is significantly negative, violating our  $\pi$ -test. In contrast, the significantly positive estimated income elasticity of 0.48 satisfies our  $w$ -test. Unlike class 2, WTP to reduce risk to all household members is significantly larger than to reduce risk to oneself (a factor of 1.4 ( $= \exp(0.339)$ )). However, WTP elicited in the parallel second valuation question for respondents in single-person households is similar (a factor 1.5 ( $= \exp(0.378)$ )), which raises the question whether the greater WTP for the household reflects scope sensitivity or a question-order effect. WTP to reduce risk to a child is 1.5 times more ( $= \exp(0.377)$ ) and to reduce risk to another adult is 1.7 times more ( $= \exp(0.522)$ ) than for oneself. Except for the coefficient on the symptom description, other coefficients on disease attributes are not statistically different from zero. Respondents who received a description of the disease were willing to pay slightly less, a factor of 0.85 ( $= \exp(-0.166)$ ), than respondents who did not receive the description.<sup>17</sup> Predicted median and mean WTP (median = 1096 €, mean = 1943 €) do not exceed the average Class 1 income, (3100 €).

In contrast to class 1, estimates for class 3 are sensitive to risk reduction. The coefficient on log risk reduction is significantly different from zero and not from one, which satisfies our  $e$ -test. WTP is insignificantly correlated with baseline risk satisfying our  $\pi$ -test. Finally, the estimated income elasticity, 1.30, is significantly positive, which satisfies our  $w$ -test. Class 3 results also show that respondents not living alone are willing to pay about 5.0 ( $= \exp(1.61)$ ) times more for a risk reduction addressed to the household than to themselves. However, as for class 1, the WTP for the second risk reduction by respondents living alone is even larger; in this case almost 7.5 times ( $= \exp(2.016)$ ). There is no significant incremental WTP to reduce risk to a child or other adult rather than to the respondent, and the other coefficients are also not statistically different from zero. Finally, predicted WTP vastly exceeds income; the predicted WTP for a mortality risk reduction of 1/10,000 is more than 200 000 €.

The LCA analysis suggests that, although the posterior estimates are coherent, there is great underlying heterogeneity.<sup>18</sup> Models of WTP for two of the three latent classes are not consistent with a standard model of preferences; only the model for Class 2, representing 42 percent of responses, is consistent with theoretical predictions. It is not surprising to find noisy answers in a self-administered internet survey. What is novel is that we are able to sort respondents by their probability of providing answers consistent with the standard model of WTP.

---

<sup>17</sup>The fact that WTP decreases with information about the morbidity of the disease is consistent with theory. In the appendix we provide an extension of our theoretical model allowing utility to depend on a continuous measure of health. We show that under reasonable assumptions a negative relationship between WTP and increased morbidity on the fatal-disease state is to be expected (Finkelstein et al. 2013).

<sup>18</sup>Andersson et al. (2016) employ LCA in analyzing responses to a discrete choice experiment valuing mortality and illness risks. They also find heterogeneity in preferences, with one of their two classes more sensitive than the other to the magnitude of the risk reduction.

Table 6: Latent Class regression

	Class (1)	Class (2)	Class (3)	Posterior
Log-risk reduction	0.128 (0.129)	0.738** (0.287)	1.559** (0.738)	0.80*** (0.339)
Log-baseline risk	-0.599* (0.326)	0.918 (0.676)	-2.639 (1.848)	-0.55 (0.744)
Log-income	0.477*** (0.0925)	0.301 (0.218)	1.306*** (0.461)	0.64 (0.514)
Child is at risk	0.377** (0.184)	1.021*** (0.361)	0.886 (1.059)	0.80*** (0.424)
Adult is at risk	0.522*** (0.143)	0.754*** (0.280)	0.202 (0.784)	0.53*** (0.221)
Household is at risk	0.339*** (0.121)	-0.248 (0.239)	1.611** (0.705)	0.46 (0.391)
Self is at risk (2de question)	0.378** (0.165)	0.387 (0.299)	2.016** (0.842)	0.86 (1.467)
Brain	0.0234 (0.128)	-0.133 (0.268)	0.960 (0.731)	0.23 (0.323)
Liver	-0.143 (0.129)	-0.464* (0.277)	0.390 (0.742)	-0.12 (0.458)
White Blood Cells	-0.00783 (0.131)	0.326 (0.256)	0.783 (0.729)	0.36 (0.589)
Latency is 10 years	0.0297 (0.112)	0.117 (0.214)	0.401 (0.646)	0.17 (0.305)
Latency is 20 years	0.107 (0.110)	-0.527** (0.228)	0.0478 (0.628)	-0.18 (0.327)
Cancer	0.00208 (0.0959)	-0.116 (0.187)	0.336 (0.504)	0.05 (0.268)
Symptom description	-0.166* (0.0989)	-0.485** (0.192)	-0.174 (0.500)	-0.30 (0.388)
Constant	-0.284 (2.849)	15.69*** (6.010)	-5.253 (16.72)	4.98 (8.921)
Sigma	1.12*** (0.041)	2.06*** (0.114)	4.07*** (0.562)	2.37*** (0.504)
Size of the class	0.29	0.42	0.29	

*Notes:* Dependent variable is WTP, measured using a double-bounded elicitation method. Follow up bids are double or halved, if the respondents agree, or disagree, to pay the initial bid. Respondents answers to WTP for each risk reduction in the study are pooled. Respondents idiosyncratic shocks are allowed to be correlated between questions. The log-risk reduction variable is takes the value of  $\log(1/10,000)$  if the respondents are faced with 1/10,000 with a risk reduction and takes the value of  $\log(2/10,000)$  if the respondents are faced with a 2/10,000 risk reduction. The log baseline risk variable takes a value of  $\log(4/10,000)$  if the baseline risk is 4/10,000 and  $\log(3/10,000)$  otherwise. Posterior standard errors are computed using parametric bootstraps (100 reps). Standard errors clustered at the respondent level are in parenthesis.

\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

Table 7 reports estimates of the marginal effects of demographic characteristics on the probability of membership in each class. The coefficients are estimated jointly with the coefficients in Table 6. The dependent variable is the posterior individual probability of belonging to a class; that is, the probability of belonging to a class conditional on the respondents' answers to the valuation questions. We include as covariates respondent demographics such as education, income, age and type of household. We include a measure of respondents' probability comprehension, defined as correctly selecting the dominant alternative on the first training question (Alberini et al. 2001). We also use time spent completing the survey as a possible proxy for cognitive effort (Nielsen et al. 2011, Rubinstein 2013). De-briefing questions such as whether respondents felt confident when answering the valuation questions, whether they expressed high concern about pesticides in food, whether they thought that if one household member became sick all would, and whether they judged their actual risks to be similar to the stated hypothetical risks are included. Finally, information about respondents' food-purchase and consumption patterns is also included.

Demographic characteristics are associated with class membership. Respondents with at least a high school education have a higher probability of membership in class 1. Higher income is associated with a higher probability of membership in class 2 and a lower probability for class 3. Age is positively associated with membership in classes 1 and 3 and negatively with class 2. Single-person households also have lower probability of class 2 membership.

Time spent completing the survey is associated with a higher probability of membership in class 1 and (less strongly) class 2, and a lower probability of membership in class 3. Respondents who answered the first training question correctly have a higher chance of belonging to class 1. Respondents believing the risk of illness is perfectly correlated among household members are less likely to belong to class 2 and more likely to belong to class 3. Respondents who considered the hypothetical risks to be similar to actual risks, and those who rarely purchase bio (organic) food were more likely to belong to class 2 and less likely to belong to class 3.

The heterogeneity in estimated WTP is summarized in Figure 3. The horizontal axis is logarithmic. The estimated density of log WTP from the LCA (bold line) reflects much greater heterogeneity than that of the standard model (fine line). The density for the LCA reveals three peaks corresponding to the three classes (from left to right, the peaks correspond to class 2, class 1, and class 3).

Table 8 presents the mean VSL and SR-VSL from class 1 and 2 estimates. Estimates for class 3 are implausibly large, so we do not report them. Estimates are for the sample-mean respondent, but these estimates are means over the error distribution, not medians as in Table 5. Because the LCA controls for much of the heterogeneity that is hidden in the error term of the standard model, the estimated residual variance and hence the difference between predicted mean and median is much smaller for the LCA than the

Table 7: Marginal effects of demographics on Class-membership

	Class 1	Class 2	Class 3
Log-income	0.028 (0.019)	0.063*** (0.024)	-0.088*** (0.022)
Education	0.1142** (0.054)	0.01 (0.059)	-0.124** (0.056)
Couples	0.05 (0.035)	-0.0365 (0.038)	-0.0135 (0.036)
Monoparental	0.031 (0.036)	-0.0612*** (0.029)	0.030 (0.039)
Family	0.0323 (0.062)	-0.1636*** (0.035)	0.131** (0.071)
Log-time spent filling survey	0.029*** (0.006)	0.0162* (0.0095)	-0.0454*** (0.010)
Age respondent	0.0015** (0.0007)	-0.0035*** (0.0004)	0.0021*** (0.0007)
Did not fail training	0.1297*** (0.052)	-0.0226 (0.062)	-0.107 (0.058)
Respondent somewhat confident	0.0331 (0.021)	-0.026 (0.024)	-0.007 (0.022)
Pesticides in food	-0.0147 (0.044)	-0.1196*** (0.041)	0.134** (0.064)
If a member gets sick, all get sick?	0.0738 (0.091)	-0.2463*** (0.073)	0.1785*** (0.076)
Considering hypothetical and actual risk are similar	0.0151 (0.035)	0.0875*** (0.033)	-0.102** (0.046)
Frequency purchase food for household	-0.036 (0.032)	-0.0213 (0.041)	0.057 (0.042)
Frequency purchase BIO	-0.0092 (0.058)	0.1795*** (0.042)	-0.1703*** (0.069)
Frequency preparing meals	-0.0685** (0.031)	0.0384 (0.034)	0.03 (0.031)
Repeated bids and risk reduction	-0.003 (0.0275)	0.0143 (0.035)	-0.0113 (0.034)

*Notes:* The horizontal sum over the three columns is equal to zero. This is due to the constraints that the probabilities must sum one. The estimates can be found in the appendix. Standard errors clustered at the respondent level are in parenthesis.

\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

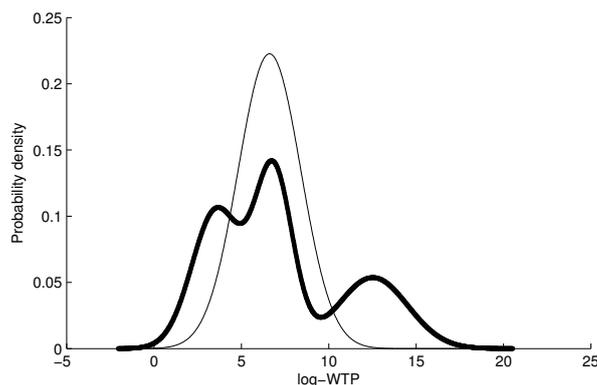


Figure 3: Log-WTP, Gaussian Mixture versus a normal distribution

standard model.<sup>19</sup>

Class 2 estimates of VSL for the respondent him or herself are approximately 2 million €. Estimates for a child and for another adult are about 6 million €. These predicted mean values are smaller than the predicted median values from the standard model and are comparable to the median SR-VSLs from the standard model (Table 5). The smaller mean estimates for class 2 than median estimates from the standard model may be due, in part, to the fact that the standard model does not downweight contributions from class 3, which satisfy most of the validity tests but have implausibly high WTP. Table 8 reports VSL and SR-VSL values when different organs are affected, and latency and disease type are changed. The only large effect is that VSL is significantly smaller when disease latency is 20 years, as compared with 1 or 10 years.

## 6 Robustness Checks

We report in Table 10 single-bounded estimates for the standard model (Cameron & Quiggin 1994). As compared with the double-bounded results presented in Table 4, we find less sensitivity to risk reduction. Qualitative results on sensitivity to baseline risk and to income do not change; the estimated incremental WTP to reduce risk to a child or other adult is smaller than for the double-bounded estimates and insignificant for the child. Estimated WTP to reduce risk to all household members jointly is significantly smaller than to reduce risk to the respondent alone, possibly reflecting an embedding effect at the household level (Kahneman et al. 1992). The single-bounded results yield a statistically larger WTP to reduce risk of cancer than of another disease, consistent with some previous literature (Van Houtven et al. 2011; Viscusi et al. 2014).

<sup>19</sup>The ratio of predicted mean to predicted median VSL in the standard model (2) is 450 ( $= \exp((3.495^2)/2)$ ) but is only 8.3 ( $= \exp((2.06^2)/2)$ ) for class 2 of the LCA. The large ratio of predicted mean to median for the standard model is consistent with many previous CV studies that report only the predicted median (for the mean respondent). For example, Hammitt & Haninger (2010) report the predicted median; their predicted mean is about 350 times larger. Hanhemann et al. (1991) advise using median VSL.

We also conducted a LCA with single-bounded estimates. The preferred model, given the Bayesian Information Criterion, has  $C = 2$  classes. The first and second class average membership probabilities are 46% and 54%, respectively. As the preferred number of classes differs between the single- and double-bounded estimates, it is difficult to compare the results directly. For the single-bounded results, presented in Table 11, WTP is sensitive to risk reduction in both classes, but the coefficient on log risk reduction is significantly different from both zero and one in both classes, violating our  $e$ -test. Class 1 has a negative and significant coefficient for baseline risk, violating our  $\pi$ -test, while Class 2 has a positive and insignificant coefficient, satisfying the  $\pi$ -test. The estimated coefficient on income is significantly positive for Class 2 and insignificantly positive for Class 1, satisfying our  $w$ -test. Neither class satisfies all three tests, but Class 2 satisfies two of three.

## 7 Conclusion

This study employs contingent valuation to elicit preferences for mortality risk reduction. We explore response heterogeneity among individuals to assess the robustness of current best-practice procedures in CVM. In particular, we explore the implications of allowing for response heterogeneity on standard validity tests, i.e., scope sensitivity, baseline risk, income, and budgetary tests. We construct a theoretical model, derive several validity tests for CV estimates of the value of small changes in mortality risks, and apply these tests to data collected from an internet panel representative of the French adult population.

Respondents were asked to value risk reductions to themselves and to other members of their households. The order of targets was randomly assigned. We find that scope sensitivity improves when we analyze only responses to the first valuation questions presented to respondents, suggesting that the consistency with validity criteria we find is not an artifact of arbitrary coherence (Ariely et al. 2003, Goldberg & Rosen 2007, Andersson et al. 2016).

A conventional analysis of our data, using a single regression model, satisfies the validity tests. However, our latent class analysis reveals substantial heterogeneity of responses that is concealed in the standard analysis. In LCA, each class represents an unequally weighted composite of the full sample of respondents and hence estimates of VSL obtained for one class are estimates based on an unequal weighting of responses from all respondents.

We find much heterogeneity in our survey responses, as illustrated by Figure 4. Each sub-figure represents the kernel density function of individual estimated posterior coefficients, and each displays multiple peaks, corresponding to the different classes. Clearly such heterogeneity, if ignored, could lead to invalid conclusions.

One of the classes identified by LCA (class 3), satisfies our primary validity tests (scope sensitivity, income sensitivity, and baseline risk insensitivity) but yields an infeasibly large predicted WTP, vastly exceeding respondent income. A possible explanation is that these responses are consistent with a contribution model, in which "responses are better de-

Table 8: Mean value per statistical life - Latent Class Analysis

	Class 1		Class 2	
	VSL	SR-VSL	VSL	SR-VSL
VSL for self	17.4 (1.99)	5.09 (1.91)	2.38 (0.48)	1.88 (0.27)
VSL for child	28.95 (5.68)	8.47 (13.77)	6.39 (2.31)	5.05 (5.38)
VSL for another adult	35.58 (4.16)	10.41 (8.31)	5.80 (1.43)	4.58 (2.24)
VSL for per HH member	11.63 (3.07)	3.40 (5.05)	0.61 (0.46)	0.49 (0.23)
VSL for self (2de question)	28.95 (4.44)	8.47 (8.77)	2.59 (0.76)	2.05 (0.61)
<b>VSL for self when:</b>				
Brain is affected	17.87 (2.81)	4.75 (3.22)	1.87 (0.51)	1.32 (0.25)
Bladder is affected	19.07 (2.92)	5.07 (3.48)	2.00 (0.52)	1.4 (0.27)
Liver is affected	14.57 (2.09)	3.87 (1.81)	1.37 (0.38)	0.96 (0.14)
White blood cells are affected	18.99 (2.64)	5.05 (2.92)	3.10 (0.70)	2.18 (0.49)
Disease latency is 1 year	16.4 (2.33)	4.36 (2.26)	2.28 (0.55)	1.61 (0.30)
Disease latency is 10 years	16.49 (2.14)	4.38 (1.96)	2.49 (0.60)	1.75 (0.36)
Disease latency is 20 years	19.85 (2.77)	5.27 (3.22)	1.41 (0.34)	0.99 (0.11)
Disease is not cancer	17.86 (2.28)	4.75 (2.23)	2.07 (0.43)	1.46 (0.20)
Disease is cancer	17.26 (2.16)	4.59 (2.01)	1.94 (0.45)	1.37 (0.21)
Symptom description provided	16.81 (2.21)	4.47 (2.09)	1.55 (0.37)	1.09 (0.14)
Symptom description not provided	25.15 (3.45)	4.88 (2.20)	3.11 (0.75)	1.82 (0.29)

*Notes:* Values are in millions of €. WTP is calculated using the specification from each model. Mean VSL is estimated for the mean individual in the following way: first, we take the exponential log-WTP for the mean individual at the average risk reduction 1.5/10000 risk reduction and we add the variance over two; second, we divide by the average risk reduction (1.5/10,000). Mean SR-VSL is estimated for the mean individual in the following way: first, we take the exponential log-WTP for the mean individual when presented with a 2/10000 risk reduction and a 1/10000, separately, while adding the variance over 2 for each; second, take the difference between the exponential log-WTP when risk reduction is equal to 2/10000 and the exponential log-WTP when risk reduction is equal to 1/10000; third, divide the result from the second step by the difference between high and low risk reduction. Standard errors are in parenthesis (delta method)

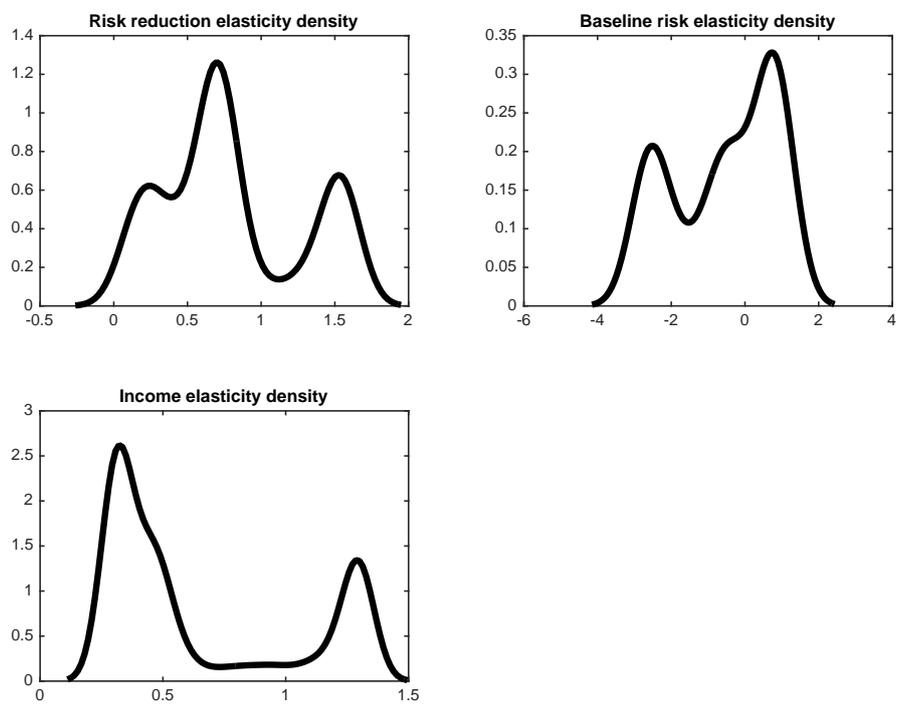


Figure 4: Scope, baseline risk and income elasticity estimated densities

Table 9: Demographics conditional on Class membership

	Class 1	Class 2	Class 3	All sample
Age of person at risk	43.22 (14.08)	41.71 (13.11)	44.41 (12.18)	42.92 (13.15)
Female	0.49 (0.50)	0.51 (0.50)	0.56 (0.50)	0.52 (0.50)
Education	0.85 (0.36)	0.80 (0.40)	0.69 (0.46)	0.78 (0.41)
Wage	3100 (1780)	2880 (1645)	2782 (1750)	2908 (1716)
Health current (VAS)	78.15 (15.25)	79.58 (15.48)	77.49 (17.82)	78.58 (16.17)
Health illness (VAS)	44.37 (27.02)	49.5 (27.62)	52.97 (28.71)	49.18 (27.95)
Health current (EQ-5D)	0.86 (0.12)	0.87 (0.13)	0.87 (0.18)	0.87 (0.14)
Health illness (EQ-5D)	0.44 (0.34)	0.52 (0.33)	0.51 (0.37)	0.50 (0.35)
N	265	435	300	1000

*Notes:* Respondents are attributed to the class where the individual conditional membership probability is highest. The means are taken over the number of respondents attributed to each class. Where female is equal to one is women, education is equal to 1 if the respondent has at least a college degree.

scribed as expressions of attitudes than as indications of economic value, contrary to the assumptions of the contingent valuation method" (Kahneman et al., 1992). Respondents who express a high degree of concern about pesticides in food have a higher probability of class 3 membership (Table 7). High concern might lead to over-reactions (Patt & Zeckhauser 2000) in the form of high WTP estimates, and not necessarily to scope insensitivity (Sunstein & Zeckhauser 2010).

As class 2 responses are consistent with theoretical validity tests, should one adopt estimates using results for this class? Sunstein (2013) argues that regulators should use preferences that are informed and rational and that "when a behavioural market failure is involved, appropriate adjustments should be made to WTP, and the VSL that emerges from WTP should be corrected accordingly." From a welfare point of view, Adler (2012) argues that preferences, which are fully informed and fully rational, should be the ones taken into account. One may ask if class 2 is representative of the French population. As shown in Table 9, there are no large differences in the distribution of demographic characteristics between the weighted sample that constitutes class 2 and the unweighted sample of respondents. Of course, this partial analysis is based on only observable characteristics. Hence, if one accepts the need to use only estimates of preferences that are arguably consistent with the standard economic model, one might use the class 2 estimates for VSL: about 2 million € for adults (own risk) and 6 million € for children.

While a finding that WTP is much smaller than proportional to risk reduction is com-

mon in CV (Hammit & Graham 1999, Carson 2012), VSL is computed and interpreted using an expected utility model, which predicts near-proportionality of WTP to risk reduction. Moreover, CV studies using the double-bounded binary choice method usually report median and not mean VSL, presumably because estimated mean VSL is implausibly large (Hammit & Haninger 2010). Moreover, median estimates of VSL are robust to outlier valuation responses, like those represented by class 3. We find that allowing for a more flexible distribution function (a mixture of normal distribution) using LCA controls for substantial heterogeneity and allows for credible mean VSL estimates. Moreover, our proposed supplementary VSL measure, SR-VSL, provides estimates under the standard analysis that are close to those found using LCA.

Our finding that WTP to reduce risk to a child or other adult in the household is larger than to reduce risk to oneself is consistent with other literature. Factors that have been identified as contributing to a divergence in WTP include age (Chanel et al. 2014, Aldy et al. 2008), risk perception (Hammit & Haninger 2010), and the perspective or context of valuation (e.g., societal or individual, altruism). Empirical studies suggest that perspective and altruism substantially influence WTP (Dickie & Ulery 2001). While the differences between values for children and adults are not problematic for benefit-cost analysis, the difference between VSL for another adult and for oneself is. Bergstrom (2006) suggests that when performing a benefit-cost analysis the sum of private valuations should be compared with the costs. This implies we should use WTP from private risk reductions (to self) to compute VSL.

While recognition of heterogeneity has proven important in other areas of economics, relatively little attention has been paid to it in the CVM literature.<sup>20</sup> Recognizing heterogeneity of responses in CVM suggests the average probability for one of our respondents to satisfy our theoretical criteria is less than half. More research is needed to evaluate the implications of accounting for heterogeneity in CVM.

---

<sup>20</sup>For example, in industrial organization introducing heterogeneity is essential when analysing consumer demands, since it allows the researcher to circumvent the independence of irrelevant alternatives assumption (IIA) implicitly introduced by the logit setting.

## References

- Adler, M., 2012. Well-being and fair distribution: beyond cost-benefit analysis. Oxford University Press.
- Alberini, A., 1995. Efficiency vs bias of willingness-to-pay estimates: bivariate and interval-data models. *Journal of Environmental Economics and Management* 29 (2), 169–180.
- Alberini, A., Hunt, A., Markandya, A., 2006. Willingness to pay to reduce mortality risks: evidence from a three-country contingent valuation study. *Environmental and Resource Economics* 33 (2), 251–264.
- Alberini, A., Krupnick, A., Cropper, M., Simon, N. B., Cook, J., 2001. The willingness to pay for mortality risk reductions: A comparison of the United States and Canada. FEEM Discussion Paper 92.01, Milan, Italy.
- Aldy, J. E., Viscusi, W. K., 2008. Adjusting the value of a statistical life for age and cohort effects. *The Review of Economics and Statistics* 90 (3), 573–581.
- Alolayan, M. A., Evans, J. S., Hammitt, J. K., 2015. Valuing mortality risk in Kuwait: Stated-preference with a new consistency test. *Environmental and Resource Economics*, published online 3 September, DOI 10.1007/s10640-015-9958-1.
- Andersson, H., Hole, A. R., Svensson, M., 2016. Valuation of small and multiple health risks: A critical analysis of sp data applied to food and water safety. *Journal of Environmental Economics and Management* 75, 41–53.
- Ariely, D., Loewenstein, G., Prelec, D., 2003. Coherent arbitrariness: Stable demand curves without stable. *Quarterly Journal of Economics*.
- Banerjee, S., Murphy, J. H., 2005. The scope test revisited. *Applied Economics Letters* 12 (10), 613–617.
- Bateman, I. J., Willis, K. G., 2001. Valuing environmental preferences: theory and practice of the contingent valuation method in the US, EU, and developing countries. Oxford University Press on Demand.
- Bergstrom, T., 2006. Benefit-Cost in a benevolent society. *American Economic Review* 96 (1), 329–353.
- Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., 2008. How are preferences revealed? *Journal of Public Economics* 92 (8), 1787–1794.
- Cameron, T. A., DeShazo, J., 2010. Differential attention to attributes in utility-theoretic choice models. *Journal of Choice Modelling* 3 (3), 73–115.
- Cameron, T. A., DeShazo, J., 2013. Demand for health risk reductions. *Journal of Environmental Economics and Management* 65 (1), 87–109.

- Carson, R. T., 2012. Contingent valuation: A practical alternative when prices aren't available. *The Journal of Economic Perspectives* 26 (4), 27–42.
- Carson, R. T., Mitchell, R. C., 1989. Using surveys to value public goods: The contingent valuation method. *Resources For the Future*.
- Chanel, O., Luchini, S., et al., 2014. Monetary values for risk of death from air pollution exposure: a context-dependent scenario with a control for intra-familial altruism. *Journal of Environmental Economics and Policy* 3 (1), 67–91.
- Chestnut, L. G., Rowe, R. D., Breffle, W. S., 2012. Economic valuation of mortality-risk reduction: Stated preference estimates from the United States and Canada. *Contemporary Economic Policy* 30 (3), 399–416.
- 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 (2), 165–184.
- Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* 39 (1), 1–38.
- Desaigues, B., Rabl, A., Ami, D., My Kene, B., Masson, S., Salomon, M.-A., Santoni, L., 2007. Monetary value of a life expectancy gain due to reduced air pollution: Lessons from a contingent valuation in France. *Revue d'Economie Politique* 117 (5), 675–698.
- Diamond, P. A., Hausman, J. A., 1994. Contingent valuation: Is some number better than no number? *The Journal of Economic Perspectives* 8 (4), 45–64.
- Dickie, M., Ulery, V. L., 2001. Valuing health in the household: Are kids worth more than parents? Tech. rep., Association of Environmental and Resource Economists 2001 Workshop: "Assessing and Managing Environmental and Public Health Risks.
- Duan, N., 1983. Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association* 78 (383), 605–610.
- EuroQol, G., 1990. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1 (1), 199–208.
- Finkelstein, A., Luttmer, E. F., Notowidigdo, M. J., 2013. What good is wealth without health? the effect of health on the marginal utility of consumption. *Journal of the European Economic Association* 11 (s1), 221–258.
- Goldberg, I., Roosen, J., 2007. Scope insensitivity in health risk reduction studies: A comparison of choice experiments and the contingent valuation method for valuing safer food. *Journal of Risk and Uncertainty* 34 (2), 123–144.

- Hammitt, J. K., 2000. Valuing mortality risk: theory and practice. *Environmental Science & Technology* 34 (8), 1396–1400.
- Hammitt, J. K., Graham, J. D., 1999. Willingness to pay for health protection: inadequate sensitivity to probability? *Journal of risk and uncertainty* 18 (1), 33–62.
- Hammitt, J. K., Haninger, K., 2007. Willingness to pay for food safety: sensitivity to duration and severity of illness. *American Journal of Agricultural Economics* 89 (5), 1170–1175.
- Hammitt, J. K., Haninger, K., 2010. Valuing fatal risks to children and adults: Effects of disease, latency, and risk aversion. *Journal of Risk and Uncertainty* 40 (1), 57–83.
- Hammitt, J. K., Liu, J.-T., 2014. Effects of disease type and latency on the value of mortality risk. *Journal of Risk and Uncertainty* 28 (1), 73–95.
- Hammitt, J. K., Robinson, L. A., 2015. The income elasticity of the value per statistical life: Transferring estimates between high and low income populations. *Journal of Benefit-Cost Analysis* 2 (1), 1–29.
- Hanemann, M., Loomis, J., Kanninen, B., 1991. Statistical efficiency of double-bounded dichotomous choice contingent valuation. *American Journal of Agricultural Economics* 73 (4), 1255–1263.
- Hausman, J., 2012. Contingent valuation: from dubious to hopeless. *The Journal of Economic Perspectives* 26 (4), 43–56.
- Hess, S., Ben-Akiva, M., Gopinath, D., Walker, J., 2011. Advantages of latent class over continuous mixture of logit models. Institute for Transport Studies, University of Leeds. Working paper.
- HM-Treasury, 2011. The green book: appraisal and evaluation in central government.
- Jones-Lee, M., 1974. The value of changes in the probability of death or injury. *The Journal of Political Economy* 82 (4), 835–849.
- Kahneman, D., Knetsch, J. L., 1992. Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management* 22 (1), 57–70.
- Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society* 47 (2), 263–291.
- Lindhjem, H., Navrud, S., Biaisque, V., Braathen, N., 2012. Mortality risk valuation in environment, health and transport policies.
- Liu, L., Neilson, W. S., 2006. Endogenous private safety investment and the willingness to pay for mortality risk reductions. *European Economic Review* 50 (8), 2063–2074.

- Machina, M. J., 1992. *Choice Under Uncertainty: Problems Solved and Unsolved*. Vol. 14. Springer Netherlands.
- McDonald, R., Chilton, S., Jones-Lee, M., Metcalf, H., 2016. Dread and latency impacts on a VSL for cancer risk reductions. *Journal of Risk and Uncertainty* 52, 137–161.
- Nielsen, J. S., Kjær, T., 2011. Does question order influence sensitivity to scope? empirical findings from a web-based contingent valuation study. *Journal of Environmental Planning and Management* 54 (3), 369–381.
- Patt, A., Zeckhauser, R., 2000. Action bias and environmental decisions. *Journal of Risk and Uncertainty* 21 (1), 45–72.
- Quinet, E., et al., 2014. L'évaluation socioéconomique des investissements publics. Tech. rep., HAL.
- Rubinstein, A., 2013. Response time and decision making: an experimental study. *Judgment and Decision Making* 8 (5), 540–551.
- Sunstein, C. R., 2013. The value of a statistical life: some clarifications and puzzles. *Journal of Benefit-Cost Analysis* 4 (2), 237–261.
- Sunstein, C. R., Zeckhauser, R., 2010. Dreadful possibilities, neglected probabilities. *The Irrational Economist, Making Decisions in a Dangerous World*.
- Train, K. E., 2009. *Discrete choice methods with simulation*. Cambridge University Press.
- Van Houtven, G., Johnson, F. R., Kilambi, V., Hauber, A. B., 2011. Eliciting benefit–risk preferences and probability-weighted utility using choice-format conjoint analysis. *Medical Decision Making* 31 (3), 469–480.
- 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 (1), 5–76.
- Viscusi, W. K., Huber, J., Bell, J., 2014. Assessing whether there is a cancer premium for the value of a statistical life. *Health economics* 23 (4), 384–396.
- Whynes, D. K., Philips, Z., Frew, E., 2005. Think of a number... any number? *Health Economics* 14 (11), 1191–1195.

## Robustness Checks Tables

Table 10: Willingness to pay results: Probit analysis

	(1)	(2)	(3)	(4)
Log-risk reduction	0.107 (0.0688)	0.178** (0.0873)	0.107 (0.0687)	0.181** (0.0872)
Log-baseline risk	0.0818 (0.162)	0.0760 (0.205)	0.0852 (0.162)	0.0770 (0.205)
Log-income	-0.0108 (0.0721)	0.00496 (0.0734)	-0.00852 (0.0723)	0.00377 (0.0736)
Child is at risk	0.233*** (0.0556)	0.195 (0.123)	0.230*** (0.0557)	0.205* (0.123)
Adult is at risk	0.174*** (0.0469)	0.245*** (0.0915)	0.171*** (0.0476)	0.247*** (0.0925)
Household is at risk	-0.185** (0.0916)	-0.168* (0.0871)	-0.195** (0.0921)	-0.171* (0.0885)
Self is at risk (2de question)	0.0252 (0.0424)	0.0291 (0.0648)	0.0173 (0.0435)	0.0264 (0.0667)
Brain			0.00825 (0.0498)	0.0169 (0.0732)
Liver			-0.0193 (0.0488)	-0.0164 (0.0738)
White Blood Cells			0.00395 (0.0498)	0.0683 (0.0756)
Latency is 10 years			0.00114 (0.0569)	-0.000906 (0.0724)
Latency is 20 years			-0.00148 (0.0576)	-0.0510 (0.0725)
Cancer			0.0906** (0.0445)	0.104* (0.0592)
Symptom description			0.00710 (0.0705)	-0.00402 (0.0725)
Constant	3.203** (1.576)	3.544* (1.892)	3.161** (1.577)	3.542* (1.893)
Initial bid	-0.222*** (0.0189)	-0.207*** (0.0199)	-0.221*** (0.0189)	-0.207*** (0.0199)
Observations	3,190	2,000	3,190	2,000

*Notes:* Dependent variable is a log-WTP, measured using a single-bounded elicitation method. Respondents answers to WTP for each risk reduction in the study are pooled. Respondents idiosyncratic shocks are allowed to be correlated between questions. The log-risk reduction variable takes the value of  $\log(1/10,000)$  if the respondents are faced with 1/10,000 with a risk reduction and takes the value of  $\log(2/10,000)$  if the respondents are faced with a 2/10,000 risk reduction. The log baseline risk variable takes a value of  $\log(4/10,000)$  if the baseline risk is 4/10,000 and  $\log(3/10,000)$  otherwise. The order of the person to which the risk reduction was addressed is random. Model (2) and (4) report results over the subset of answers to the first valuation question, and questions on household level risk including second questions to single person households. The household WTP question is always asked last. Standard errors clustered at the respondent level are in parenthesis.

\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

Table 11: Latent Class regression: Probit

	Class 1	Class 2
Initial bid	-0.395*** (0.0427)	-0.420*** (0.0294)
Log-risk reduction	0.297** (0.143)	0.256* (0.143)
Log-baseline risk	-0.629* (0.350)	0.470 (0.332)
Log-income	0.0199 (0.0907)	0.174* (0.0983)
Child is at risk	-0.0564 (0.224)	0.439** (0.196)
Adult is at risk	-0.0835 (0.165)	0.658*** (0.135)
Household is at risk	-0.222* (0.135)	0.0377 (0.122)
Self is at risk (2de question)	-0.187 (0.168)	-0.389** (0.186)
Symptom description	0.0190 (0.0961)	-0.169* (0.0942)
Brain	0.182 (0.133)	-0.0767 (0.131)
Liver	-0.112 (0.136)	-0.0376 (0.129)
White Blood Cells	0.254* (0.147)	0.0121 (0.135)
Latency is 10 years	0.0995 (0.123)	0.0412 (0.109)
Latency is 20 years	-0.0204 (0.128)	-0.242** (0.114)
Cancer	0.204** (0.100)	0.0992 (0.0934)
Constant	1.560 (3.111)	6.349** (3.075)
Size of the class	0.46	0.54

*Notes:* Dependent variable is a log-WTP, measured using a single-bounded elicitation method. Respondents answers to WTP for each risk reduction in the study are pooled. Respondents idiosyncratic shocks are allowed to be correlated between questions. The log-risk reduction variable is takes the value of  $\log(1/10,000)$  if the respondents are faced with 1/10,000 with a risk reduction and takes the value of  $\log(2/10,000)$  if the respondents are faced with a 2/10,000 risk reduction. The log baseline risk variable takes a value of  $\log(4/10,000)$  if the baseline risk is 4/10,000 and  $\log(3/10,000)$  otherwise. Posterior standard errors are computed using parametric bootstraps (100 reps). Standard errors clustered at the respondent level are in parenthesis.

\* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

## Appendix

### Elasticity of willingness to pay with respect to risk reduction.

By differentiating  $P$  with respect to  $e$ . It follows that:

$$\frac{\partial P}{\partial e} = \frac{u_A(w_e) - u_D(w_e)}{(\pi + e)u'_A(w_e) + (1 - \pi - e)u'_D(w - P(e, w, \pi))} > 0.$$

Note that when  $e \rightarrow 0$  we have  $\frac{\partial P}{\partial e} \equiv VSL$ , and  $P = 0$ . If we multiply by  $e$  and divide by  $P$  we have:

$$\eta_e^P = \frac{u_A(w_e) - u_D(w_e)}{(\pi + e)u'_A(w_e) + (1 - \pi - e)u'_D(w - P(e, w, \pi))} \frac{e}{P(e, w, \pi)}.$$

Here  $\eta_e^P$  denotes the elasticity of substitution between the risk reduction,  $e$ , and willingness to pay,  $P$ . As we are interest in cases when  $e \rightarrow 0$ , applying l'Hôpital's rule yields:

$$\lim_{e \rightarrow 0} \eta_e^P = \lim_{e \rightarrow 0} \frac{-e \frac{\partial P}{\partial e} (u'_A(w_e) - u'_D(w_e)) + u_A(w_e) - u_D(w_e)}{\frac{\partial P}{\partial e} EU'(w_e) + P [u'_A(w_e) - u'_D(w_e) - \frac{\partial P}{\partial e} EU''(w_e)]},$$

and given that  $P(0, w, \pi) = 0$ , we find,

$$\lim_{e \rightarrow 0} \eta_e^P = \frac{u_A(w) - u_D(w)}{EU'(w) \frac{u_A(w) - u_D(w)}{EU'(w)}} = 1.$$

The relationship between willingness to pay,  $P$ , and risk reduction  $e$ , when  $e \rightarrow 0$  is that of proportionality.

### Elasticity of baseline risk on willingness to pay.

Next, we investigate the functional relationship between baseline mortality probability and willingness to pay. Thus, we differentiate with respect to  $1 - \pi$  and obtain:

$$\frac{\partial P}{\partial(1 - \pi)} = \frac{u_A(w_e) - u_D(w_e) - u_A(w) - u_D(w)}{(\pi + e)u'_A(w_e) + (1 - \pi - e)u'_D(w - P(e, w, \pi))} > 0.$$

Then, multiplying by  $1 - \pi$  and dividing by  $P(e, w, \pi)$  yields:

$$\eta_{1-\pi}^P = \frac{u_A(w_e) - u_D(w_e) - u_A(w) - u_D(w)}{(\pi + e)u'_A(w_e) + (1 - \pi - e)u'_D(w - P(e, w, \pi))} \frac{1 - \pi}{P}.$$

Here,  $\eta_{1-\pi}^P$  denotes the elasticity of substitution between baseline probability of death and willingness to pay. As we are interest in cases when  $e \rightarrow 0$ , as we did previously, applying l'Hôpital's rule yields:

$$\lim_{e \rightarrow 0} \eta_{1-\pi}^P = \lim_{e \rightarrow 0} \frac{-\frac{\partial P}{\partial e} (u'_A(w_e) - u'_D(w_e)) (1 - \pi)}{\frac{\partial P}{\partial e} EU'(w_e) + P [u'_A(w_e) - u'_D(w_e) - \frac{\partial P}{\partial e} EU''(w_e)]}.$$

Given that  $P(0, w, \pi) = 0$ , we find that:

$$\frac{1 - \pi}{\pi} \geq \lim_{e \rightarrow 0} \eta_{1-\pi}^P = \frac{1 - \pi}{\pi + \frac{u'_A(w)}{u'_A(w) - u'_D(w)} - 1} > 0.$$

Provided that the probability of survival is close to 1, the elasticity of substitution between the baseline risk of death and willingness to pay is positive but close to 0.

### Elasticity of income on willingness to pay.

We investigate the relationship between wealth and willingness to pay. Differentiating  $P$  with respect to  $w$  yields

$$\frac{\partial P}{\partial w} = 1 - \frac{\pi u'_A(w) + (1 - \pi) u'_D(w)}{(\pi + e) u'_A(w_e) + (1 - \pi - e) u'_D(w_e)}.$$

Then, multiplying by  $w$  and dividing by  $P(e, w, \pi)$  yields:

$$\eta_w^P = \frac{EU'(w_e) - EU'(w) w}{EU'(w_e) P}.$$

Here,  $\eta_w^P$  denotes the elasticity of substitution between wealth and willingness to pay. As we are interested in cases when  $e \rightarrow 0$  applying l'Hôpital's rule yields:

$$\lim_{e \rightarrow 0} \eta_w^P = \lim_{e \rightarrow 0} w \frac{u'_A(w_e) - u'_D(w_e) - EU''(w_e) \frac{\partial P}{\partial e}}{\frac{\partial P}{\partial e} EU'(w_e) + P [u'_A(w_e) - u'_D(w_e)]}.$$

which in turns yields,

$$\lim_{e \rightarrow 0} \eta_w^P = w \frac{u'_A - u'_D}{u_A - u_D} - w \frac{EU''(w)}{EU'(w)} > 0.$$

Additionally, consider the following expression of VSL:

$$VSL = \frac{u_A(w) - u_D(w)}{(\pi) u'_A(w) + (1 - \pi) u'_D(w)} > 0. \quad (16)$$

The income elasticity of VSL can be calculated by differentiating equation (16) with respect to  $w$ , dividing the result by  $VSL/w$ , and substituting the right-hand side of the equation for VSL. The income elasticity  $\eta_w^{VSL}$  can be expressed, after some simplifications as:

$$\eta_w^{VSL} = w \frac{u'_A - u'_D}{u_A - u_D} - w \frac{EU''(w)}{EU'(w)},$$

This implies that  $\lim_{e \rightarrow 0} \eta_w^P = \eta_w^{VSL}$ .

## Willingness to pay for mortality risk reduction and health

As reported in table 3, describing the disease has a negative impact on the reported health state when ill. In this subsection we explore how willingness to pay for a mortality risk reduction varies with health.

As before, we assume a one period state dependent expected utility framework. Let an individual derive utility  $u(w, h_j)$  from wealth,  $w$  and health  $h_j$ , where  $j = A, D$  denote the two possible states, alive or dead, respectively. The utility of wealth conditional on death is associated with bequest motives. We denote first (second) derivatives with respect to wealth by the subscript 1 (11) and those with respect to health by the subscript 2 (22). We adopt the standard assumptions  $u(w, h_A) > u(w, h_D)$  and  $u_1(w, h_j) > 0$ ,  $u_2(w, h_j) > 0$ ,  $u_{11}(w, h_j) \leq 0$ ,  $u_{22}(w, h_j) \leq 0$  for  $j = \{A, D\}$ . Without loss of generality, we will measure health quality on a unit scale so that  $h_A = 1$ ,  $h_D = 1 - z$ , where  $z < 1$ .

Letting  $\pi$  denote the probability of survival, expected utility is given by  $E(U) = \pi u(w, 1) + (1 - \pi) u(w, h_D)$ . Willingness to pay to reduce the risk by the amount,  $e$ , denoted  $P_z = P(e, w, \pi, z)$ , is defined by:

$$(\pi + e) u(w - P_z, 1) + (1 - \pi - e) u(w - P_z, h_D) = \pi u(w, 1) + (1 - \pi) u(w, h_D). \quad (17)$$

Taking the derivative with respect to  $z$  in equation (17) yields,

$$\frac{\partial P(e, w, \pi, z)}{\partial z} = \frac{(1 - \pi) u_2(w, h_D) - (1 - \pi - e) u_2(w - P_z, h_D)}{(\pi + e) u_1(w - P_z, 1) + (1 - \pi - e) u_1(w - P_z, h_D)}. \quad (18)$$

The sign of equation (18) depends on the effect of wealth on the marginal utility of health (i.e.,  $u_{12}$ ). If we observe WTP for mortality risk decreases with increased morbidity in the death health state, we infer that the marginal utility of health increases with lower wealth, we refer to it as positive state dependence (Finkelstein et al. 2013). Positive state dependence is a sufficient but not necessary condition for the sign of equation (18) to be negative. The following relationship need to holds:

$$\frac{u_2(w - P_z, h_D)}{u_2(w, h_D)} > \frac{1 - \pi}{1 - \pi - e}. \quad (19)$$

Whereas if we observe WTP for mortality risk reduction rise with increased morbidity in the death health state, we infer a negative state dependence and the sign of equation (19) is reversed.

## Symptom Descriptions

There are symptom descriptions for each organ, which are identical for cancer and other disease. Respondents were randomly allocated between receiving a description (for all valuation questions) or not. The symptom descriptions follow.

### Brain Cancer/Degenerative Neurological Disease

**Description:** Initially, the symptoms will be mild, but they will gradually get worse over time. [You/He/She] will have headaches that last for an hour or two at a time but the pain will be relieved with over-the-counter medicines like aspirin. At times, [your/his/her] speech will become slurred and [you/he/she] will be difficult to understand. [You/He/She] will also have difficulty with [your/his/her] arms and legs. They will feel weak and [you/he/she] will need to sit down and take a rest from what [you are/he is/she is] doing. Over time, these symptoms will get worse and may persist for most of the day. For example, [you/he/she] may wake up with severe headaches that last all day, have difficulty swallowing, and may vomit when eating or drinking. Eventually, [you/he/she] will become so weak that [you/he/she] will be confined to bed or a wheel chair most of the time. [You/He/She] will not have control over [your/his/her] normal bodily functions and will not be able to go to the bathroom or shower by [yourself/himself/herself]. [You/He/She] may start to have seizures that will cause [you/him/her] to lose consciousness and make [your/his/her] body shake uncontrollably. Once the seizures become difficult to control, [you/he/she] will die within a few months.

### Bladder Cancer/Bladder Disease

**Description:** Initially, the symptoms will be mild, but they will gradually get worse over time. [You/He/She] will have blood in your urine that appears red or brown. The blood may be present some days but not others, and may disappear for weeks at a time. Sometimes blood clots will form that will cause pain or obstruction to the flow of urine. [You/He/She] will have pain when trying to urinate. [You/He/She] and will feel that [your/his/her] bladder has not completely emptied and that [you need/he needs/she needs] to urinate frequently, even when [you are/he is/she is] unable to do so. Over time, these symptoms will get worse and may persist for most of the day. Eventually, [you/he/she] will lose a significant amount of [your/his/her] muscle mass and will become so weak that [you/he/she] will be confined to bed or a wheel chair most of the time. [You/He/She] will not have control over [your/his/her] normal bodily functions and will not be able to go to the bathroom or shower by [yourself/himself/herself]. After reaching this point, [you/he/she] will die within a few months.

### **Liver Cancer/Liver Failure**

**Description:** Initially, [you/he/she] may feel normal and go about [your/his/her] daily activities with little difficulty, but at some point [your/his/her] skin will turn yellow and [you/he/she] will experience discomfort in [your/his/her] upper belly that may last for hours at a time. [You/He/She] may experience bleeding, bruising, and swelling from fluid in the belly and legs. Eventually, [you/he/she] will have a poor appetite and feel tired much of the time. Over time, these symptoms will get worse and may persist for most of the day. [You/He/She] will become so weak that [you/he/she] will be confined to bed or a wheel chair most of the time. [You/He/She] will not have control over [your/his/her] normal bodily functions and will not be able to go to the bathroom or shower by [yourself/himself/herself]. [You/He/She] will eventually have periods of time when [you/he/she] will be confused and unable to think clearly because toxins normally cleared by the liver will build up in [your/his/her] body. [You/He/She] may experience insomnia and daytime drowsiness, confusion, inability to recognize people and places, and finally, coma. Once this occurs, [you/he/she] will die within a few months.

### **Lymphoma /White Blood Cell Disorder**

**Description:** Initially, the symptoms will be mild, but at some point [you/he/she] will feel very weak and sick with high fevers, severe sweats, and shaking chills. [You/He/She] will wake up with the sheets drenched. [You/He/She] may have pain in [your/his/her] abdomen that lasts for hours at a time. [You/He/She] will have repeated infections and will need to go to the hospital for intravenous medication. Over time, these symptoms will get worse and eventually [you/he/she] will become so weak that [you/he/she] will be confined to bed or a wheel chair most of the time. [You/He/She] will not have control over [your/his/her] normal bodily functions and will not be able to go to the bathroom or shower by [yourself/himself/herself]. At some point, medication will not be able to fight the infections [you get/he gets/she gets]. [You/He/She] may pass in and out of consciousness, experience shortness of breath, and finally, coma. Once this occurs, [you/he/she] will die within a few months.

## EM-algorithm

The EM-algorithm is iterative. EM exploits the fact that, although the class membership does not depend on the choices made, the choices provide information about the class membership. Suppose that an agent is vegetarian, but we do not know. Observing her food choices consecutively would lead us to infer, with a high degree of certainty, that she is a vegetarian. The key part of EM algorithms is updating the belief of an individual membership in a class  $s$ , which is done through Bayes theorem. Let  $h_{is}(x_{it}|y_{it})$  be individual  $i$ 's posterior probability of belonging to class  $s$ . It is computed as follows:

$$h_{is}(x_{it}|y_{it}) = \frac{L_i^s}{\sum_{c=1}^C L_i^c}. \quad (20)$$

Note that  $L_{is}$  corresponds to individual  $i$ 's contribution to the overall likelihood, which is given by the sequence of answers,  $\prod_{t=1}^T Q_{it}(\theta_s, x_{1it}, y_{it})$ , conditional on being a class  $s$  type of individual, weighted by the probability of being a member of class  $s$ ,  $\pi_{is}(x_{2it}, \alpha_s)$ . Given the evidence (her observed choices), we update our beliefs on individual  $i$ 's membership by weighting her contribution to the likelihood on each of the distinct classes  $C$ . If the contribution to a class, say  $s_1$ , is higher than the others, then it would be reflected in our higher posterior beliefs,  $h_{is_1}(x_{it}|y_{it})$ .

From an empirical point of view, estimating (15) is computationally complex. An alternative log-likelihood,  $\mathcal{E}(\theta)$ , can be maximized to yield the same parameters (Train 2008). It is defined as follows:

$$\mathcal{E}(\Theta) = \sum_{i=1}^N \sum_{c=1}^C h_{is}(x_{it}|y_{it}) \log(L_i^c).$$

Note that  $\log(L_i^c)$  can be sub-divided into two parts:

$$\log(L_i^c) = \log\left(\prod_{t=1}^T Q_{it}(\theta_s, x_{1it}, y_{it})\right) + \log(\pi_{is}(x_{2it}, \alpha_s)).$$

The log-likelihood is then given by:

$$\mathcal{E}(\Theta) = \sum_{i=1}^N \sum_{c=1}^C h_{is}(x_{it}|y_{it}) \log\left(\prod_{t=1}^T Q_{it}(\theta_s, x_{1it}, y_{it})\right) + \sum_{i=1}^N \sum_{c=1}^C h_{is}(x_{it}|y_{it}) \log(\pi_{is}(x_{2it}, \alpha_s)), \quad (21)$$

where the first term in the RHS of equation (21) will be named  $LL_\theta$  and the second term will be  $LL_\alpha$ . Moreover, since  $\sum_{c=1}^C \pi_{ic}(x_{2it}, \alpha_c) = 1$  we will assume that:

$$\pi_{is}(x_{2it}, \alpha_s) = \frac{\exp(\alpha_s x_{2it})}{\sum_{c=1}^C \exp(\alpha_c x_{2it})}, \quad (22)$$

and we impose the following identification restriction,  $\alpha_C = 0$ , so that the coefficients from each class are interpreted with respect to class  $C$ .

As noted earlier, the model has to be estimated in an iterative fashion. We build the algorithm in Matlab. The algorithm is as follows:

1. Form the contribution to the likelihood  $L_i^s$  for each class  $s = 1, \dots, C$ .
2. Form the individual-specific posterior probabilities of class membership  $h_{is}^r(x_{it}|y_{it})$ , where  $r$  denotes the  $r^{th}$  iteration.
3. Maximize each class-specific WTP regression  $LL_\theta$  to obtain the updated sets of  $\theta_s^{r+1}$  with  $s = 1, \dots, C$ . Each regression uses as weights the posterior probabilities of class membership computed in step 2.
4. Maximize jointly the prior probability logit functions  $LL_\alpha$  to obtain the updated sets of  $\alpha_s^{r+1}$  with  $s = 1, \dots, C - 1$ . Each prior is weighted by the posterior probabilities of class membership computed in step 2.
5. Repeat step 1 to 4 until convergence.

Although it is simple, the EM algorithm is quite slow to converge (Train 2009) and it can converge to a local maximum. We used several starting points and set the change in the log-likelihood function  $LL(\Theta)$  to be smaller than  $1e^{-10}$  to ensure convergence to a global maximum.<sup>21</sup>

---

<sup>21</sup>Note that there are more sophisticated variants of the EM algorithm (simulated annealing, stochastic EM), which tend to be more robust to being trapped in local optima.