Believe only half of what you see

the role of preference heterogeneity in contingent valuation

James K. Hammitt*
Daniel Herrera-Araujo†

April 30, 2015

Abstract

The paper shows, under an expected-utility framework, the exact theoretical relationship between willingness to pay (WTP) to reduce small mortality risks, risk reduction, baseline risk, and income. We propose a scope-revealing value per statistical life (SR-VSL) that accounts for any lack of scope sensitivity. Using a French stated-preference survey fielded to a large, nationally representative internet panel, we explore by how much, and why, respondents depart from the expected utility predictions. We find that only 40% of our respondents' behave as predicted by expected-utility theory. High concern for environmental risks to health, low education, and less time spent completing the survey are good predictors of deviant answers. Our preferred value per statistical life estimates range from 2.2 to 3.4 million € for adults, and over 6 million € for children. No differences are found for disease-specific WTP, particularly, we find no evidence of a premium for cancer.

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, Phone: 617.432.4343, E-mail: jkh@harvard.edu

†INRA, Lerna, Toulouse School of Economics, 21 Allée de Brienne 31000 Toulouse, France, E-mail: daniel.herrera@tse-fr.eu

1 Introduction

"From dubious to hopeless", this famous title from Hausman (2012), crystallizes the opposition that hypothetical markets, or the contingent valuation method (CVM) has generated in the literature. Despite the criticism, CVM remains one of the main sources for estimating the marginal rate of substitution between small changes in the probability of death and wealth, or "Value per Statistical Life" (VSL).[1] "Is some number better than no number?" Diamond et al. (1994) argue that if CVMs do not correctly elicit preferences the answer is no. It is crucial to have, assess and report theoretical validity checks of correctly elicited preferences if the results from a CVM are to be used for policy.

How to assess validity, when the goal is to elicit respondents preferences over small changes in mortality risks? Provided respondents see the results as potentially influencing governments, and care about the outcome of the survey, economists can use theoretical predictions to assess validity (Carson & Groves 2007). Theoretically (Carson & Mitchell 2013), respondents' answers should be sensitive to characteristics that matter. In our context, respondents willingness to pay (WTP) should increase with the scope of the good. Specifically, WTP should increase with the magnitude of the risk reduction. For small risk reductions, WTP should increase near-proportionally to its size (Corso et al., 2001). In addition, respondents' WTP to reduce a risk to an entire household should be at least as large as the minimum WTP to reduce that risk to any individual member. Additionally, respondents' WTP should be sensitive to wealth, specifically, in our context, WTP should not decrease with wealth. Moreover, respondents answers should not be sensitive to characteristics that, in theory, do not matter, such as question framing. Specific to our context, respondents WTP should not vary with small differences in baseline risks, (Graham & Hammitt 1999; Hammitt 2000a).[2]

We present stated-preference estimates of WTP to reduce mortality risks to identified individuals: the respondent him or herself, their child, another adult living in the household, or the household. Hypothetical mortality risks are associated with pesticide residues on food, and risk reductions are embodied through an alternative food produced following a hypothetical, "pesticide security system" warranted by the state. Risks are described as a function of baseline risk of illness (with the conventional food type), risk reduction (with the alternative food type), affected organ (brain, bladder, liver, lymphocytes), disease type (cancer, non-cancer), and latency period (1, 10, 20 years). These characteristics are randomly varied across respondents using a full factorial design. Estimates are obtained using a representative French internet panel, CSA. A total of 1000 respondents completed our survey.

[1]VSL accounts for the lion's share of benefits in many cost-benefit assessments. A retrospective analysis of the Clean Air Act indicates that mortality risks account for 95 percent of the present value of monetized benefits from 1970 to 1990 (EPA 1997; Hammitt & Robinson 2011). As a result, different values of VSL may radically change the spectrum of alternative policies that could be cost-beneficial.

[2]As shall be seen in section 2, by theoretical validity we assume that the relevant model is that of an agent that maximizes its expected utility. Expected utility is the canonical theory of choice under uncertainty in economics (Jones-Lee, 1974).

This article has the following objectives: (1) to propose an additional measure of VSL that reflects the quality of respondents' answers; (2) to implement our theoretical predictions and to investigate by how much, and why, individuals depart from the expected utility framework, by looking at respondents' heterogeneity in preferences; (3) To provide new estimates of VSL for the French population and how these vary depending on characteristics of the disease and affected individual.

Despite the measure's policy relevance, only a few papers have tried to estimate VSL in France.[3] Most of them have issues regarding their economic validity.[4] We propose VSL estimates for France that satisfy economic validity criteria. Our estimates range from 2.2 to 3.4 million € for adults, and 6 million € for children. No differences are found for disease specific WTP, particularly, we find no evidence of higher WTP to reduce risk of cancer compared to other fatal diseases.

Our survey instrument is adapted from one administered in the United States by Hammitt and Haninger (2010). In the US, it produced results consistent with our validity criteria, i.e., WTP increasing an nearly proportional to magnitude of the risk reduction, independent of small differences in baseline risk, and increasing with income. In our French sample, responses are also broadly consistent with validity criteria. However, looking at individual preference heterogeneity casts doubts over the apparent validity of our CVM. Although respondents' answers are consistent with theoretical predictions, when they are endogenously classified into three or more homogeneous subgroups (using latent class analysis), a rich underlying story is revealed. In total, 59 % of our sample violates the predictions derived from the standard expected-utility model. The data reveal that 30% of respondents have a willingness to pay that greatly exceeds their monthly income, thus possibly violating their budgetary restriction. While, 29% of our respondents have a WTP that increases with a lower baseline risk, violating the insensitivity to baseline risk argument. We find that what drives the membership to the remaining 41% is, spending more time completing the survey, as well as expressing less concern about environmental quality than other respondents.

The study is organized as follows: Section II provides a theoretical background and a comprehensive literature review on CV validity; Section III provides details on the survey design; Section IV reports the econometric model, Section V reports results, Section VI discusses the results, and concludes.

---

[3] In 2013 the French administration updated its guidelines for project evaluation with the Rapport Quinet. The VSL endorsed by the Rapport was extracted from a recent OECD meta-analysis done in 2012, which contained all available studies eliciting VSL in France. Most estimate monetary values for a risk reduction associated with transportation, or pollution. Only one focuses on valuing risk reductions for various forms of cancer or other type of degenerative diseases (Oken et al. 2012).

[4] For example: using the same questionnaire as Alberini et al. (2006), Desaigues et. al (2007) estimated the valuation of life expectancy gain due to a reduction of air pollution in France. They report a large embedding effect (Kahneman et al. 1992) between the risk reduction questions of 1 and 5 in 1000. They report a ratio between the WTP for 5/1000 and 1/1000 of 1.6; theory suggests that it should have been close to 5. The estimates for the value of a life year (VOLY) range from 0.02 to 0.22 million € with a mean VSL of 4.12 million €.

2 What makes a CV study credible?

Carson & Groves (2007)[5] identify two backbone criteria and a property that have to be satisfied if the survey is to produce policy relevant information: (1) respondents need to believe the survey could influence government actions; (2) respondents must care about the outcome of those actions. Carson & Groves (2007) term a survey satisfying both criteria as consequential. [6] They argue that only consequential surveys can be interpreted in economic terms. Furthermore, they put forward a *face-value property*, which they define as "the property that respondents always truthfully answer the specific survey question being asked. There are two aspects of this property: (a) that respondents always answer truthfully, and (b) that respondents always correctly understand and answer the question being asked."

First, we will set out our theoretical background and then we will discuss how some of our theoretical implication have been dealt with in the literature.

2.1 Theoretical background

We assume a one period, state dependent expected utility framework to explore the monetary trade-off that consumers face when considering a reduction in risk. Take a simple preference specification where utility derives from wealth (w), $u_j = u_j(w)$, where $j = A, D$ denote the two possible states, alive or dead, respectively. The utility of death is associated with bequest motives.

Assuming π denotes the probability of survival, expected utility is given by $E(\mathcal{U}) = \pi u_A(w) + (1-\pi) u_D(w)$. Let the willingness to pay to reduce the risk by the amount, e, denoted by $P(e, w, \pi)$, be defined by:[7]

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

where u_j is such that $u_j' > 0$ and $u_j'' \leq 0$ for $j = \{A, D\}$. Moreover, we assume that the utility of income is larger when alive than dead, $u_A > u_D$, as well as for marginal utility of income, $u_A' > u_D' \geq 0$.[8] Note that when $e = 0$ then $P(e, w, \pi) = 0$.

Moreover, the marginal rate of substitution between risk reduction, e and wealth, w is then:

[5] Their research question can be summarized by: "does a survey question need to meet certain conditions before it can be expected to produce useful information about an agent's (respondent's) preferences?"

[6] Here are examples of inconsequential questions: "(a) being asked of a population or at a location that is irrelevant from the perspective of an agency seeking input on a decision, (b) providing few, if any, details about the goods and how they would actually be provided, (c) asking about goods that are implausible to provide, or (d) about an implausible prices for them." Carson & Groves (2007).

[7] We assume that both, e and π, are exogenous to the individual.

[8] If we assume the utility of bequest to be zero, the willingness to pay, $P(e, w, \pi)$, we can re-express the equation above in the following way: $P(e, w, \pi) = w - u_A^{-1}\left(\frac{\pi}{\pi + e} u_A(w)\right)$ Here, it is clear that $P(e, w, \pi) < w$.

$$\frac{\partial P\left(e,w,\pi\right)}{\partial e}=\frac{u_A\left(w-P\left(e,w,\pi\right)\right)-u_D\left(w-P\left(e,w,\pi\right)\right)}{\left(\pi+e\right)u_A^{'}\left(w-P\left(e,w,\pi\right)\right)+\left(1-\pi-e\right)u_D^{'}\left(w-P\left(e,w,\pi\right)\right)}>0. \quad (1)$$

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

$$VSL=\frac{\partial P\left(0,w,\pi\right)}{\partial e}\equiv\frac{\partial P_0}{\partial e} \quad (2)$$

Let, η_e^{wtp}, η_w^{wtp}, and $\eta_{1-\pi}^{wtp}$, denote the elasticity of substitution between willingness to pay $P\left(e,w,\pi\right)$ and the risk reduction e, income w, and baseline risk probability, 1-π, respectively. Moreover denote by η_w^{VSL}, the elasticity of substitution between VSL and income. The following results hold:

$$\lim_{e\to 0}\eta_e^{wtp}=1 \quad (3)$$

$$\eta_w^{VSL}=\lim_{e\to 0}\eta_w^{wtp}>0 \quad (4)$$

$$\frac{1-\pi}{\pi}\geq\lim_{e\to 0}\eta_{1-\pi}^{wtp}=\frac{1-\pi}{\pi+\frac{u_A^{'}(w)}{u_A^{'}(w)-u_D^{'}(w)}-1}>0 \quad (5)$$

$$P\left(e,w,\pi\right)<w. \quad (6)$$

Hence, for any utility function satisfying our assumptions, if the risk reduction is small enough, an increase of the risk reduction by 1% increases willingness to pay by 1%, an increase of income increases by a positive percentage willingness to pay, and an increase of baseline risk has virtually no effect on WTP.[9]

If we accept in the standard expected-utility model,[10] equations (3), (4), (5), and (6), provide powerful, yet simple, testable implications, which are key in assessing whether a contingent valuation survey is valid.

Table 1 summarizes the empirical tests to be performed. As we deal with household level risk reductions we provide two additional tests corresponding to our specific context. Each can be regarded as a form of scope sensitivity test: (1) WTP for a risk reduction affecting everyone in a household should be at least as large as the minimum WTP for any individual person living in the household; (2) when households are composed of a single individual, the differences between WTP to reduce risk to the household and to the individual WTP should be zero.

[9] The reader can find the derivation of the three functional relationships in the appendix.

[10] Hammitt (2000a), makes a parallel between the standard expected-utility model and alternative theories of decision making under uncertainty. He argues that the sole requirement to satisfy near-proportionality is local linearity in probabilities (Machina 1992). One case where near-proportionality does not hold is when willingness to pay functions are not smooth in the risk reduction (Kahneman & Tversky 1979). For example, when individuals are willing to pay for a risk reduction only if the risk reduction is above a certain threshold that they consider meaningful, the results above do not hold.

Table 1: Validity tests summary

Characteristics	Criterion	Test	Name of the test
Risk reduction	$\eta_e^{wtp} = 1$	$\hat{\beta}_1 > 0$,$\hat{\beta}_1$= 1, s.	RR-test
Baseline risk	$\eta_{1-\pi}^{wtp} \approx 0$	$\hat{\beta}_2 = 0$, n.s.	BLR-test
Income	$\eta_w^{wtp} > 0$	$\hat{\beta}_3 > 0$, s.	INC-test
Budget	$w > P(e, w, \pi)$	$\hat{P} <$ Income	B-test
Risk reduction for not single person HH	HH-WTP $\geq$ min ind. WTP	$\beta_6 \geq \min\{0, \beta_4, \beta_5\}$	HH-WTP1
Risk reduction for single person HH	HH-WTP = WTP for self	$\beta_7 = 0$, n.s.	HH-WTP2

Notes: To fix ideas, consider a case in which we observe WTP, $P(e, w_i, \pi)$, for a risk reduction e. Let WTP be defined as: $log(P_i(e, w_i, \pi)) = \beta_1 log(RR_i) + \beta_2 log(BLR_i) + \beta_3 log(INC_i) + \beta_4 Child_i + \beta_5 OAdult_i + \beta_6 HH_i + \beta_7 SPH_i + z_i \beta_8 + \xi + \epsilon_i$, where, RR_i, BLR_i and INC_i correspond to the risk reduction, baseline risk and income variables, respectively. $Child_i$, $OAdult_i$, SPH_i and HH_i correspond valuations of risk reductions addressed to a child, on other adult, a single person household or a household composed of more than one individual.

s. and n.s. denote significant and not significantly different from zero, respectively. $\hat{P}$ denotes the estimated willingness to pay for the mortality risk reduction. HH-WTP denotes willingness to pay to reduce risk addressed to the entire household.

2.2 Scope sensitivity in the literature

The prior section describes theoretical predictions and validity tests for CVM valuing small risk reductions. We now discuss how some of these predictions have been faced in the literature.

There is strong opposition against the validity of CVM. The main issue raised is the embedding effect. It occurs if "the same good is assigned a lower value if WTP for it is inferred from WTP for a more inclusive good rather than if the particular good is evaluated on its own" (Kahneman & Knestch 1992, p. 58). The persistence of the embedding effect across CVMs in time, and settings, makes it the most worrisome issue in the CVM literature. The failure of sensitivity to scope in CVM has been interpreted, (1) as a failure of CVM as measurement tool to elicit preferences (Kahneman & Knetsch 1992; Diamond et al. 1994), and (2) as reflecting the incapacity of individuals to form preferences over (public) goods (Diamond et al. 1994).

Alternatively, others have considered that the failure might be due to a poor survey design. Hammitt & Graham (1999) report that only 9 out of 25 CVM studies of reductions in health risks are found to exhibit scope sensitivity. They conclude that addressing the scope insensitivity issue is a question of respondents' correct understanding of the good being valued, requiring a good 'study design'. [11] Corso et al. (2001) find that the use of visual aids reduces the scope insensitivity problem. Moreover, Corso et al. (2001)

[11] Consumers tend to commit mistakes when low probabilities are at hand, even in real world situations. Citing Carson (2012): "low-probability risks are often poorly understood in contingent valuation surveys, as they are by consumers in the real-world behaviour involving financial planning and insurance decisions." (p.35)

distinguish between strong and weak scope sensitivity, where strong refers to WTP that is nearly proportional to risk reduction and weak to WTP that is statistically significantly increasing with risk reduction. From the 9 studies identified by Hammitt & Graham (1999), none exhibited strong scope sensitivity.

Understanding the trade-off between risk reduction and wealth is closely correlated with cognitive effort exerted during the task. Time spent on answering the survey, as a proxy for cognitive effort, might explain the weak scope sensitivity. Nielsen et al. (2011) analyse the relationship between scope sensitivity and response time. They find a negative relationship between scope sensitivity and time spent on completing the questionnaire: the quicker respondents answer, the lower the probability of being sensitive to scope.

More recently, Rubinstein (2013) adopts the fast and slow perspective advocated by Kahneman (2011). The idea suggests two types of reasoning: (1) a fast and instinctive one, dubbed system 1, and (2) a slow cognitive one, system 2. His finding suggests that quick, and instinctive, respondents are more prone to error than those who take their time. He highlights that response time, although noisy, is a useful tool for the evaluation of experimental results. Visual aids would be useless if the respondents are simply clicking away the survey. For this reason, we include this type of para-data into our analysis.[12]

Weak, and even lack of, scope sensitivity could also find its roots in individuals' limited attention resources. Cameron & DeShazo (2013) implement a structural model accounting for individuals' attention limitation. Individuals may care about many attributes but given their constrained attention, they can attend to only a limited set. When a model yields an estimated marginal utility of zero, it may be misleading; it is not that individuals do not care about the characteristic, it is simply too costly to attend to it. Our approach mirrors Cameron and DeShazo (2013) model from a reduced form point of view. We are able to identify groups of respondents who attend to some characteristics, but not others.

In some cases, when non-satiation is violated, scope sensitivity might not be a necessary condition for a CV to elicit preferences. Banerjee et al. (2005) provides a simple example to show that scope is not necessary: "[...] consider a consumer whose preferences are represented by a utility function given by $U(a, b) = \min\{a + b, 2b\}$. The expansion path of this utility lies along a 45° line through the origin; a typical indifference curve is piecewise linear with slope -1 above the 45° line and slope zero below. Pick any bundle, (a, b) where a lies on or below the 45° line. Since the indifference curve through the bundle (a, b) is horizontal for any increment B of a, scope sensitivity is violated. But because the preferences of this consumer are represented by a utility function, her preferences are regular. Hence regular preferences do not guarantee scope." (p.6). [13] Moreover, scope

[12] Para-data are the data generated by the respondents while completing the survey. They concern how respondents answered, not what they answered. The time the respondents took to complete the survey and the number of clicks the respondents made are examples of para-data. This valuable source of information is obtainable through the use of internet surveys, and has been largely unexploited in the contingent valuation literature.

[13] Banerjee et al. (2005) shows that only under assumptions of continuous, strongly monotonic and total preferences we should expect scope sensitivity. It is only under these assumptions that the validity of the scope sensitivity test can hold.

sensitivity is not a sufficient condition for validity; e.g. if WTP exceeds individual wealth.

2.3 Baseline Risk in the literature

Under our standard model, WTP is insensitive to small changes in baseline risk. This does not hold under other modelling assumptions. For example, allowing agents to self-protect, by introducing a risk reducing technology, Liu et al. (2006) showed that the relationship between the baseline probability of death and willingness to pay to reduce fatality risks could be negative. Also, Breyer et al. (2002) find that when bequest motives, along with a sufficient amount of non-inheritable capital are allowed, the relationship between baseline risk and WTP is negative. Despite providing models where a negative relationship between WTP and baseline risk can survive, neither provide a sense of the magnitude of the effect.

Finally, empirical evidence of a negative relationship can be found in Smith & Desvousges, (1987). They estimated WTP to pay to reduce risk of death from hazardous waste and found a negative relationship between WTP and baseline risk.

2.4 Income elasticity in the literature

There is a general consensus in the theoretical literature that income elasticity of VSL is positive. Eeckhoudt & Hammitt (2001), as well as Kaplow (2005) derive that, under an expected utility framework, the relative risk aversion coefficient for wealth is a lower bound for income elasticity of VSL. Hence if an agent is risk averse, her income elasticity should be positive. The connection can be understood as follows. VSL depends on the marginal utility cost of expenditures to reduce mortality risks. It follows that VSL depends on how the marginal utility cost of such expenditures varies with income levels, in other words, how the marginal utility of income falls as income increases. The coefficient of risk aversion is the measure of this rate (Evans & Smith 2010).

Arrow's (1971) seminal work on behaviour under uncertainty suggests that the coefficient of risk aversion should be at least 1. (Watt and Vasquez 2012). Empirical estimates are in the order of 1, 10 or above (Kaplow 2005, Campbell 2003, Chetty 2003).

Evans & Smith (2010) construct a theoretical setting that, unlike Kaplow (2005), introduces behavioural changes to exogenous income shocks, e.g., a spouse enters the labour market if the other spouse faces unemployment. As a consequence, income elasticity would be smaller than with no behavioural changes. Moreover, by allowing consumption and labour to be complementary, the elasticity is also decreased; a higher level of consumption would decrease the dis-utility for an additional hour of work. Both results suggest that income elasticity might be smaller than the coefficient of relative risk aversion.

Empirically, income elasticity of VSL can be estimated from, at least, two sources: (1) wage-differential studies, (2) contingent valuation. For the former, Viscusi & Aldy (2003) survey a relevant body of the literature and find that income elasticity of VSL is in the range of 0.5 and 0.6 with the upper bound of the 95 percent confidence interval falling below 1. More recently, a study by Doucouliagos et al. (2014) finds income elasticity estimates to be between 0.25 and 0.6; the study includes both wage-differential studies, as well as stated-preference estimates. [14] The latter method also provides valid estimates for income elasticity. Alberini et al. (2004) find elasticities ranging from 0.2 to 0.3 on a multi-country study. Similarly, Hammitt & Haninger (2010) find elasticities of 0.1 to 0.3 in a study of valuing pesticide risks to adults and children .

Notwithstanding, contingent valuation studies do not always find a significant relationship between income and VSL. Does that imply an automatic rejection of validity? Given theoretical and empirical evidence, we believe that verifying income elasticity is at least not negative is a viable validity check.

2.5 How to compute VSL, empirically?

Implicitly CVMs valuing small risk reduction assume, when computing the VSL, that respondents behave as expected utility theory predicts. Deviations from the canonical expected utility framework, reflected by scope insensitivity, need to be accounted for.

When eliciting preferences, generally, the respondents are asked to consider discrete hypothetical mortality risk reductions. It follows that there are several statistics a researcher can use to approximate the VSL. To fix ideas, let e_c denote a risk reduction between, $e_1 = \frac{1}{10000}$ and $e_2 = \frac{2}{10000}$. Furthermore, let P_{e_j} denote the willingness to pay for risk reduction, e_j, where $j = \{1, 2, c\}$. With this setting, there are at least three ways of computing a value per statistical life:[15]

$$VSL_{e_1} = \frac{P_{e_1}}{e_1}. \tag{7}$$

where VSL_{e_1} corresponds to the VSL obtained when proposing a risk reduction of e_1. Or,

$$VSL_{e_2} = \frac{P_{e_2}}{e_2}. \tag{8}$$

where VSL_{e_2} corresponds to the VSL obtained when proposing a risk reduction of e_2. Finally,

$$VSL_{e_c} = \frac{P_{e_2} - P_{e_1}}{e_2 - e_1} = \eta_e^{wtp} \frac{P_{e_1}}{e_1} = \frac{\partial P_{e_c}}{\partial e} \tag{9}$$

where VSL_{e_c} corresponds the value of statistical for a risk reduction e_c and η_e^{wtp} corresponds to the WTP elasticity with respect to the risk reduction. We will denote VSL_{e_c} as SR-VSL. In the literature, the most prevalent statistics used are (7) and (8), and often

[14] Why does the size of the income-elasticity matter? A policy-oriented reason is transferability. Countries that do not have reliable estimates of VSL use income elasticity estimates to extrapolate another country's reliable VSL to theirs. For example, the Quinet Report (2013) (baseline document for economic policy evaluation in France) has extrapolated the VSL values found in the 2012 OECD meta-analysis to obtain the French VSL. The income elasticity used for the calibration amounts to 0.8%.

[15] Please refer to the appendix to have a detailed explanation of how we derive them.

the mean between both. To our knowledge, (9) is not used.

Theoretically, the differences between these measures are minimal for small risk reductions (in the limit as $e_2 \to 0$ they are equal). Empirically, a large majority of contingent valuation studies addressing mortality risk reductions suffer from lack of scope sensitivity, which leads the statistics to differ substantially. Despite the latter issue, the use of statistics (7) and (8) provide researchers/policy-makers with positive, statistically significant, and perhaps misleading values (Diamond et al. 1994).

Regarding the quality of CVM, economic theory suggests a dichotomous approach; either it is good or not. If there is appropriate scope-sensitivity, VSL computed from (9) is very close from to the value calculated (7), whereas, if there is no scope-sensitivity the value is close to zero. In this respect, the use of (9) reveals information about the scope sensitivity of the CVM. For the previous extreme cases, it suggests that SR-VSL is a better choice than either VSL_{e1}, or VSL_{e2}. For weak scope sensitivity it is not as clear cut.[16] What is clear is that, under weak scope sensitivity, SR-VSL is the lower bound reflecting the quality of the survey with lower values for VSL.

3 Survey design

3.1 Structure of the questionnaire and survey administration

The survey was conducted in 2012, with the goal to elicit WTP for a reduction in the probability of death from consuming pesticide residues on food. The questionnaire was identical, save for language and other minor differences, to the questionnaire used by Hammitt & Haninger (2010). The survey was administered to a random sample of the CSA internet panel. Panel members were recruited through random e-mails and closely matched to the french national population with quotas on age, socio-economic factors, gender and geographical variables. Data were gathered in 2 waves between July and August 2012. We had 1000 completed interviews.

Respondents were asked to value reductions in the risk of a fatal disease that might affect a specified target: himself or herself, a child (aged between 2 and 18 years) or another adult living in their household. The risk was described as due to pesticide residues on food that only the individual would eat. Reduction of the risk was made possible 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 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 the complementary probability of no illness were distinctively coloured. The adverse health effect was described as a chronic fatal disease, either cancer or non-cancer, affecting the

[16] If scope sensitivity is 0.5, when theoretically we expect 1, does it mean that the quality of the survey is 50%?

bladder, brain, liver or blood. The symptoms of the disease would first appear after a latency period of 1, 10, or 20 years. Respondents were asked to evaluate the current health of the target individual, and their health conditional on suffering the specified illness, using a numerical scale on which 100 corresponds to full health and 0 to a state as bad as dead, and using the EQ-5D health state classification system (reference).

Before the valuation questions, respondents were presented with two practice questions with feedback. In the first, one food type was both safer and less expensive than the other. Respondents who chose the dominant alternative were told that the food they had selected was both safer and less expensive than the other and that this was the logical choice. Respondents who chose the dominated alternative were told that the food they had selected was both less safe and more expensive than the other and invited to choose again. In the second practice question, neither alternative was dominant. Respondents were told the food 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.

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 he or she 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. 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.

A total of 1000 respondents and 2263 risk reductions are included in the analysis. Our final sample consists on 186 single-person households, 284 households that include at least one other adult and no child (younger than 18 years old), 125 households that include no other adult and at least one child, and 359 households that include at least one other adult and one child. Non-response to questions regarding monthly household income was of about 16%. The missing values were imputed as the average answer conditional on the type of household.

3.2 Data

Table (2) reports on demographics. Sample means and standard deviations are taken for the entire sample, and for each sub-sample 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 44 years with a fifty percent chance that the respondent is a female and has a bachelors degree. The net monthly income (in 2012 €) of the average household is of 2885 €, and they are rather concerned by the quality of the environment.

Table (3) reports on survey para-data and design. Median time to complete the survey

Table 2: Household summary statistics

	Pooled	Self	Child	Other Adult
Age of person at risk	35.87	44.10	9.25	44.71
	(19.12)	(12.82)	(5.15)	(14.20)
Female	0.53	0.57	0.47	0.51
	(0.49)	(0.50)	(0.50)	(0.50)
Current health state	81.08	78.39	91.97	78.81
	(12.91)	(16.26)	(8.70)	(17.36)
Disease health state	48.20	48.03	46.03	51.97
	(26.35)	(27.63)	(30.50)	(29.84)
Current health (EQ-5D score)	0.88	0.86	0.95	0.87
	(0.10)	(0.14)	(0.10)	(0.14)
Disease health state (EQ-5D score)	0.48	0.48	0.45	0.51
	(0.33)	(0.34)	(0.37)	(0.34)
Loss in EQ-5D score when ill	0.40	0.38	0.51	0.36
	(0.33)	(0.34)	(0.38)	(0.36)
% questions asked is not first	0.76			
	(0.42)			
Environmental concern	3.71			
	1.14			
Income	2884			
	(1675)			

Notes: Female is a dummy variable taking the value 1 when female, 0 otherwise. 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. Environmental concern is a self reported variable ranging from 1(low) to 5(high). Income corresponds to household net income in 2012 euros.

of 17 minutes. The means for baseline risk, risk reduction, latency, cancer and affected organ confirm that randomization was successful.

Table 3: Para-data summary statistics

	Pooled	S.D.
Time completing the survey	17.23	7.14
Baseline risk	3.5	0.50
Risk Reduction	1.5	0.50
Latency	10.4	7.78
Cancer	0.48	0.50
Bladder	0.24	0.43
Brain	0.25	0.43
Blood	0.24	0.43
Sample size	1000	

Notes: Time completing the survey corresponds to the median time. Baseline risk and risk reduction are per 10000 persons. Latency is over 10, 20 or 30 years. Cancer, is a dummy variable equal to 1 if the disease is described as cancer. Bladder, Brain, Blood, are equal to 1 if the affect organ is bladder, brain or blood, respectively, zero otherwise. The omitted organ is the liver.

4 Empirical implementation

Unobserved individual heterogeneity abounds in contingent valuation studies: individuals differ in their cognitive resources andmay differ in the set of characteristics to which they attend (Cameron & DeShazo 2013). Understanding such heterogeneity is key.

4.1 Identification

A key point in the identification of preferences for risk reductions is the assumption that respondents care about risk reductions and think that the study outcome influence government decision making.

Assumption 1: Respondents would demand the risk reduction at no cost.

The survey is designed in such a way that respondents might be able to relate to the question being asked, detailed information is provided regarding the good being valued and plausible provisions of the good and prices are given.

Assumption 2: Respondents do not have kinked preferences.

Assumption 2 is necessary to interpret scope insensitivity as a failure in understanding the good being valued rather than identification of kinked-preferences.

Variation in bids, disease characteristics and target individuals allow the predicted probabilities to vary, which generates enough moments to identify the coefficients. Finally, the panel structure helps in the identification of respondent classes (Greene 2008).

4.2 Estimation

Latent Class Regressions (LCR) is a valuable method to assess such unobserved heterogeneity (Train 2008). In a recent paper, Hess et al. (2011) suggest Latent Class models as being able to retrieve richer patterns of heterogeneity than continuously mixed models. We assume that the underlying coefficients follow a discrete distribution, and LCR non-parametrically estimates such distribution; coefficients, as well as their weights. Hence, we are able to group together individuals that have similar preferences.

Assume that there are N agents, who report their WTP in T choice occasions. Following our theoretical model, define the observed WTP, $P(e, w, \pi)$, of respondent i which belongs to class s, where $s = \{1, \ldots, C\}$ and C the number of classes, in choice occasion t for a risk reduction e as:

$$log(P_{it}(e, w_i, \pi)) = \beta_{1s} log(RR_{it}) + \beta_{2s} log(BLR_{it}) + \beta_{3s} log(INC_i) + z_{it}\beta_{4s} + \xi_s + \epsilon_{ist} \quad (10)$$

where RR_{it}, BLR_{it} and INC_i correspond to risk reduction, baseline risk, and income, respectively; z_{it} contain other individual characteristics, including targeted individual dummies (child, one other adult and household); ξ_s correspond to a constant unobservable class s fixed effect and ϵ_{ist} rationalizes all remaining choice-to-choice individual variation.

The agents are assumed to know their WTP for a risk reduction, e, but the value is not observed by econometricians. A double-bounded method is used to determine agents' WTP up to an interval (Hanemann et al. 1994). A first bid is proposed to the agents, which they can decide to accept to pay or not. A second question follows, where the initial bid is halved or doubled depending on the respondents' initial answer; if "yes" the initial bid is doubled, if "no" the bid is halved.

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 affect WTP, x_{1it}, and characteristics that explain membership to a particular class, x_{2it}, which may or may not overlap.

We assume ϵ_{ist} follows a log-normal distribution. Hence, the conditional probability of individual i to belong to a particular WTP interval is given by:

$$Q_{it}\left(\theta_s, x_{1it}, y_{it}\right) = \begin{cases} \Phi\left(\frac{b_{itL} - x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it} = 0 \\ \Phi\left(\frac{b_{it0} - x_{1it}\beta_s}{\sigma_s}\right) - \Phi\left(\frac{b_{itL} - x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it} = 1 \\ \Phi\left(\frac{b_{itU} - x_{1it}\beta_s}{\sigma_s}\right) - \Phi\left(\frac{b_{it0} - x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it} = 2 \\ 1 - \Phi\left(\frac{b_{itU} - x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it} = 3 \end{cases} \tag{11}$$

where, Φ is the normal cumulative distribution function, $\theta_s = (\beta_s, \sigma_s)$ are the mean and standard error parameters of the normal distribution for the class s. The indicator of the choice y_{it} represent "No-No", "No-Yes", "Yes-No" and "Yes-Yes", respectively. However since θ_s is unknown, the sequence of observed choices has to be evaluated over all the possible values that θ_s can take. We assume that the density of the parameters is described by a discrete distribution. It follows that the log-likelihood function is:

$$LL\left(\Theta\right) = \sum_{i=1}^{N} \log\left(\sum_{s=1}^{C} \pi_{is}\left(x_{2it}, \alpha_s\right) \prod_{t=1}^{T} Q_{it}\left(\theta_s, x_{1it}, y_{it}\right)\right) \tag{12}$$

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

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

where,

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

The main identifying assumption is that respondents' unobserved shocks are independent between respondents and choice occasions (Train 2008). 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 which class each agents belongs to, we would have to estimate C conventional likelihoods.

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 $h_{is}\left(x_{it}|y_{it}\right)$ be individual i's posterior probability of belonging to class s. It is computed as follows:

$$h_{is}\left(x_{it}|y_{it}\right) = \frac{L_i^s}{\sum_{c=1}^{C} L_i^c}. \tag{13}$$

Note that L_{is} corresponds to individual i's contribution to the overall likelihood, which is given by the sequence of answers, $\prod_{t=1}^{T} Q_{it}\left(\theta_s, x_{1it}, y_{it}\right)$, conditional on being a class s type of individual, weighted by the probability of being a member of class s, $\pi_{is}\left(x_{2it}, \alpha_s\right)$. 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, $h_{is_1}\left(x_{it}|y_{it}\right)$.

From an empirical point of view, estimating (12) is computationally complex. An alternative log-likelihood, $\mathcal{E}\left(\theta\right)$, can be maximized to yield the same parameters (Train 2008). It is defined as follows:

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

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

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

The log-likelihood is then given by:

$$\mathcal{E}\left(\Theta\right) = \sum_{i=1}^{N}\sum_{c=1}^{C} h_{is}\left(x_{it}|y_{it}\right)\log\left(\prod_{t=1}^{T} Q_{it}\left(\theta_s, x_{1it}, y_{it}\right)\right) + \sum_{i=1}^{N}\sum_{c=1}^{C} h_{is}\left(x_{it}|y_{it}\right)\log\left(\pi_{is}\left(x_{2it}, \alpha_s\right)\right) \tag{14}$$

where the first term in the RHS of equation (14) will be named LL_θ and the second term will be LL_α. Moreover, since $\sum_{c=1}^{C} \pi_{ic}\left(x_{2it}, \alpha_c\right) = 1$ we will assume that:

$$\pi_{is}\left(x_{2it}, \alpha_s\right) = \frac{\exp\left(\alpha_s x_{2it}\right)}{\sum_{c=1}^{C}\exp\left(\alpha_c x_{2it}\right)} \tag{15}$$

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[17] The algorithm is as follows:

1. Form the contribution to the likelihood L_i^s for each class $s = 1, \ldots, C$.
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_θ to obtain the updated sets of θ_s^{r+1} with $s = 1, \ldots, C$. Each regression uses as weights the posterior probabilities of class membership computed in step 2.
4. Maximize jointly the prior probability logit functions LL_α to obtain the updated sets of α_s^{r+1} with $s = 1, \ldots, C-1$. 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 2008) 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. [18]

5 Results

The following section reports on results from a standard WTP regression analysis, as if only one class existed, and a latent class regression analysis.

5.1 Standard analysis

Respondents' willingness to pay is assumed to follow a log-normal distribution (equation 10). The coefficients of all the models presented below are estimated using maximum likelihood estimation (Alberini 1995). The standard errors are calculated using Wald test (Train 2008). We allow for correlation between answer-specific idiosyncratic errors for each respondent, but assume independence between respondents. Our sample consists of 1000 respondents and 3190 answers.

There are two types of households: Households with only one person, and households with more than one person. All respondents are asked to report their WTP to reduce risks to, when possible, three of the members in the household. Moreover, each respondent is asked to report their WTP to reduce a risk to the everyone in the household simultaneously. This questions is always asked last. The same logic applies to a single person household. A respondent from a single person household is asked about her WTP for a personal risk

[17]We modified Patrick P. C. Tsui's Matlab Code to adapt it to our needs.

[18]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.

reduction, and then she is asked about her WTP for a risk reduction addressed to the entire household, which by definition is herself. (Recall that the risk reduction, disease characteristics and bid amount differ between questions.)

Model (1) in Table (4) examines the effects of, risk reduction, baseline risk and income elasticity on WTP. As can be observed, only the coefficient on log risk reduction is significantly different from zero, but also different from 1, partially violating our RR-test. Individuals are willing to pay 1.35 ($= \exp(0.437 \log(2))$) times more for a risk reduction of 2 in 10,000 than a risk reduction of 1 in 10,000. The point estimate of log-baseline risk is not significantly different from zero, consistent with our validity check, BLR-test, but not significant on income thus failing to satisfy our income validity test, INC-test. Additionally, when asked about a risk reduction addressing all household members, respondents living in a multi-person household are willing to pay 1.5 ($= \exp(0.41)$) times more for a risk reduction to all members of the household (including themselves) than to themselves alone. For the respondents that live alone, WTP to reduce risk to the household is not significantly different than to reduce risk to themselves. Both results are consistent with our HH-WTP1 and HH-WTP2 tests. Finally, respondents are willing to pay on average 2.6 ($= \exp(0.98)$) times more to avoid a risk to their child than to themselves and 2 ($= \exp(0.68)$) times more to avoid a risk to another adult in their household.

*****

Model (2) in Table (4), was estimated over the subset of answers to the first valuation question provided by the respondents, as well as questions which concerned risk reductions addressed to all members in the household jointly. The coefficient on log risk reduction is significantly different from zero, but not from 1, satisfying our RR-test. Respondents are willing to pay 1.61 ($= \exp(0.69 \log(2))$) times more for a risk reduction of 2 in 10,000 than for a risk reduction of 1 in 10,000. As with model (1), BLR-test is satisfied while the INC-test is not. The coefficient on log-baseline risk is not significantly different from zero, nor is the coefficient on income. Finally, both HH-WTP tests are satisfied. A respondent living in a multi-person household is willing to pay to avoid a risk to all members 1.92 ($= \exp(0.66)$) times more than a risk addressed to themselves. For households with only one individual, there is no significant difference. Moreover, respondents are willing to pay 2.62 ($= \exp(0.96)$ and 2.5 ($= \exp(0.92)$) times more for a risk reduction to their child, and to another adult in their household than to themselves, respectively.

Models (3) and Model (4) include the same variables and observations as model (1) and model (2), respectively. In addition, they include characteristics on the disease assigned to each individual. Coefficients relevant to our validity criteria are not affected by adding these variable. With the exception of he coefficient on "organ affected is the liver" , in model (3), all the coefficients are insignificantly different from zero. This implies that respondents are not willing to pay more (or less) if the disease is cancer as compared to not cancer, or if the latency is 10 or 20 years as compared to 1 year.

Finally, table (5) reports on the implied VSL for model (1) and (2), and for each type of individual addressed. Both VSL and SR-VSL are computed for the mean respondent

Table 4: Willingness to pay results: Standard analysis

	(1)	(2)	(3)	(4)
Log-risk reduction	0.437**	0.690**	0.433**	0.699**
	(0.22)	(0.28)	(0.22)	(0.28)
Log-baseline risk	0.184	0.192	0.196	0.208
	(1.96)	(1.51)	(1.94)	(1.49)
Log-income	-0.027	0.015	-0.022	0.013
	(0.25)	(0.25)	(0.24)	(0.26)
Child is at risk	0.974***	0.964**	0.945***	0.948**
	(0.17)	(0.38)	(0.17)	(0.38)
Adult is at risk	0.689***	0.925***	0.643***	0.882***
	(0.14)	(0.29)	(0.14)	(0.29)
Houshold at risk is multi-person	0.408***	0.656***	0.353***	0.610***
	(0.12)	(0.20)	(0.13)	(0.21)
Houshold at risk is one person	-0.204	0.059	-0.228	0.041
	(0.28)	(0.27)	(0.29)	(0.27)
Cancer			0.198	0.175
			(0.14)	(0.19)
Brain			-0.029	-0.062
			(0.16)	(0.26)
Liver			-0.286*	-0.403
			(0.16)	(0.26)
White Blood Cells			-0.104	0.021
			(0.16)	(0.26)
Latency is 10 years			0.028	-0.084
			(0.18)	(0.23)
Latency is 20 years			0.024	-0.109
			(0.18)	(0.23)
Constant	12.14**	13.87**	12.19**	14.19**
	(5.06)	(6.19)	(5.06)	(6.20)
Sigma	3.39***	3.49***	3.39***	3.49***
	(0.15)	(0.16)	(0.15)	(0.16)
Observations	3190	2000	3190	2000

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 assumed to be independent 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. As the order of the person to which the risk reduction was addressed is random the "not first question" takes the value of 1 if the corresponding question is not the first the respondent had to answer. Model (2) and (4) report results when exluding the notfirst questions, except household questions. The household WTP question is always asked last. Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

in our sample.[19] There are two implicit assumptions underlying expressions (7) and (8): first, that the risk is close to zero; second, that there is perfect scope sensitivity. In the case of (9), the risk reduction, e_c, is small enough, as it is bounded above by $e_2 - e_1$. We relax the second assumption (perfect scope sensitivity) by taking the empirical estimate of the elasticity. As a result, the implied value per statistical life is adjusted by the scope sensitivity. In fact, if strong scope sensitivity is empirically verified (ie. near-proportionality), then equations (9) and (7) yield equivalent results. In the opposite case, if there is lack of scope sensitivity the implied value per statistical life, computed from (9) will tend to zero, while the value per statistical life computed from (7) will not be changed.

Table 5: Median value per statistical life

	Model (1)		Model (2)		Model (2)	LCA Class 2	
	VSL	SR-VSL	VSL	SR-VSL	Mean/Median	VSL	SR-VSL
VSL Self	6.33	2.76	4.45	3.11	448	0.26	0.24
	(0.95)	(1.47)	(0.95)	(1.45)		(0.06)	(0.09)
VSL Child	16.75	7.31	11.67	8.15	448	0.72	0.66
	(3.33)	(4.02)	(4.11)	(4.54)		(0.27)	(0.34)
VSL Other adult	12.61	5.5	11.23	7.83	448	0.66	0.61
	(2.19)	(2.97)	(2.58)	(3.74)		(0.18)	(0.26)
VSL per H. M.	3.17	1.38	2.86	2.01	448	0.07	0.06
	(0.54)	(0.74)	(0.53)	(0.91)		(0.02)	(0.03)
VSL Self, S. P. H.	5.16	2.25	4.73	3.3	448	0.36	0.33
	(1.59)	(1.37)	(1.50)	(1.75)		(0.11)	(0.15)

Notes: H.M. stands for Household member. S. P. H. stands for single person household. Values are in millions of euros. WTP is calculated using the specification from each model. VSL is estimated for the mean individual in the following way: first, we take the exponential WTP for the mean individual; second, we reduction with respect to the WTP for each model. Standard errors are in parenthesis (delta method). The mean VSL is computed by adding variance over two before taking the exponential. Only model's 2 mean/median ratio is reported.

It is clear from table (5) that VSL and SR-VSL do not coincide. For adults, the median value per statistical life is between 3 and 7 million Euro, while the median scope-revealing value per statistical life is around 2-5 million euros. For children, the median of value per statistical life is around 16 Million euro while, SR-VSL hovers around 7 Million Euro. Note that both model (1) and (2) have SR-VSL which are statistically identical. Large standard errors on the SR-VSL reflect the quality of elasticity of substitution. Finally, mean and median differ by a factor of 448, which may be come as a result of the functional form assumed; it allows for infinitely large values.

5.2 Latent class analysis

Assuming log-normality is simple and provides consistent estimates (Hanemann et al. 1991). Regardless, CVs are stated preferences and they rely onconsistency with the

[19] Section 3.2 provides the summary statistics for the mean respondent.

expected-utility framework to justify their validity. It follows that the average effect (or coefficient) could hide respondents who are not taking the survey seriously, or simply do not fully understand it.

We propose a latent class analysis (LCA) to better understand the underlying heterogeneity. The added value of performing a LCA is the explicit modelling of class-membership. In addition, each respondent has a has a positive probability of membership in each class. We will refer to 'Class X' members as a weighted contribution of inputs from all the respondents - with more weight given to those with high posterior probability of being in Class X.

We include education, income and environmental concern levels as class covariates. Education of respondents serves as a proxy for cognitive resources, while income serves a proxy for opportunity cost of time. Environmental concern serves as a proxy for general interest in the survey. We use time spent completing the survey as another proxy for cognitive effort (Nielsen et al. 2010, Rubinstein 2013). Finally, respondents' probability comprehension, measured by the success of the training program (Alberini et al. 2002), is also included.[20]

We perform a LCA for $C = 2, \ldots, 5$ classes. The preferred model, given the Bayesian Information Criterion, is the model with $C = 3$ classes. The first, second and third class have average posterior membership probabilities of 29% ,41% and 30%, respectively. Table (6) reports results for the 3-class model estimation, as well as the posterior estimates. The regression includes a full set of interactions with a dummy variable distinguishing the valuation question asked first from those asked later (not reported).

First, consider the posterior coefficients. The estimated model has the same covariates as those found in Table (4), model (2), and can be compared directly with the posterior coefficients of our LCA model (Train 2008). 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 the log-risk reduction is statistically different from zero and not different from 1, which satisfies our RR-test criteria. Given an increase of 1% of the risk reduction, WTP increases by 1%. The relationship between baseline risk and WTP is negative although not significant, which is consistent with our BLR-test. Income elasticity of WTP is positive, but not significantly different from zero satisfying our INC-test. Despite the non-significance of the coefficients, the point estimates of the remaining coefficients are not far from what is usually found in the literature. Noisy estimates of posterior probabilities are to be expected (Train 2008). Confidence intervals for the coefficients of the posterior model generally include the from the model (2) in Table (4); the posterior coefficients have not added any additional insight.

Next, consider the coefficients for each class-specific regression. Class 1 coefficients show that respondents are scope insensitive, violating our RR-test. Additionally, Class 1 respondents have a negative relationship between baseline risk and WTP, violating our

[20]The inclusion of more or fewer variables has no impact on the formation of the classes; they serve as explanatory variables.

Table 6: Latent Class regression

	Class 1	**Class 2**	**Class 3**	**Posterior**
Log-risk reduction	0.182	0.918***	1.911***	1.002**
	(0.13)	(0.31)	(0.72)	(0.42)
Log-baseline risk	-0.735**	0.576	-1.081	-0.301
	(0.33)	(0.72)	(1.69)	(0.94)
Log-income	0.545***	0.342	1.694***	0.806
	(0.09)	(0.23)	(0.45)	(0.58)
Child is at risk	0.463**	1.004**	0.732	0.766
	(0.18)	(0.39)	(1.05)	(1.05)
Adult is at risk	0.637***	0.924***	-0.122	0.527
	(0.14)	(0.29)	(0.72)	(0.47)
Houshold at risk is multi-person	0.523***	-0.290	1.582**	0.507
	(0.12)	(0.26)	(0.63)	(0.55)
Houshold at risk is one person	0.426**	0.329	1.50**	0.70
	(0.16)	(0.31)	(0.73)	(0.75)
Constant	4.27***	8.95***	16.46**	7.252*
	(1.25)	(3.25)	(7.00)	(4.27)
Sigma	1.02***	2.07***	3.88***	2.309***
	(0.51)	(0.15)	(0.31)	(0.36)
Size of the Class	0.29	0.41	0.30	

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 assumed to be independent 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). Robust standard errors in parenthesis.
* significant at 10%; ** significant at 5%; *** significant at 1%

BLR-test. These results do not agree with theoretical predictions from the standard model in section 4. Nevertheless, respondents in Class 1 satisfy our INC-test by having a positive income elasticity estimated as 0.55. Class 1 respondents in multi-person households are willing to pay about 1.7 times as much to reduce risk to their household as to themselves, satisfying one part of our HH-WTP test. However, single person households are willing to pay significantly more (1.5 times) to reduce risk to their households than to themselves, violating the other part. [21] Respondents are willing to pay about the same amount (between, 1.5 and 1.9 times more) for a risk reduction for a children as for another adult in their household. As a consequence of multiple violations of our validity tests, we do not consider Class 1 as a good class to base further analysis.

In contrast to Class 1, respondents in Class 3 are sensitive to risk reductions. The coefficient suggests a more than proportional relationship between risk reduction and WTP; a risk reduction of x increases willingness to pay by 1.8 x ($= \exp(1.91 \log(2)) = 1.77$), but it is not statistically different from 1. Class 3 respondents satisfy our RR-test. Respondents' WTP is insignificantly negatively correlated with baseline risk satisfying our BLR-test. Finally, Class 3 respondents satisfy our INC-test because respondents in this class have a positive and statistically significant income elasticity (1.7). When respondents live in a multi-person household, they are willing to pay almost 5 (exp (1.58)= 4.85) times more for a risk reduction addressed to the household than to themselves. Whereas when respondents live alone, their WTP is also 5 times higher for a risk addressed to the household than to a risk addressed to themselves. Only one out of our two HH-WTP are satisfied. Class 3's WTP to avoid a risk to children, and other adults, does not differ significantly from WTP to avoid a risk addressed to themselves. Finally, median (or mean) WTP exceeds median (or mean) income; the WTP is over 200 000 € for a risk reduction of 1 in 10,000. Even if we do not know the average wealth in our sample, we are confident that Class 3 respondents are not revealing their preferences for risk reductions.

Respondents in Class 2 have a point estimate on log-risk reduction of 0.9 and it is statistically different from zero and not from 1, which suggest near-proportionality. The log-baseline risk coefficient and log-income coefficient are both positive but not significantly different from zero.[22] [23] Median (or average) WTP (median = 30 €, mean = 244 €) does not exceed the average Class 2 income, (2865 €). WTP to reduce a risk addressed to the entire household, regardless of whether it is a single person household or not, is not statistically different from WTP for a risk reduction addressed to themselves. Moreover, Class 2 has WTP for risk reductions addressed to children, and other adults, 2.7 ($= \exp(1)$) times higher than WTP for a risk reduction addressed to themselves. Respondents in Class 2 satisfy all the criteria for CV validity.

[21] Though not reported in Table (6), coefficients do not significantly vary between first, and the subsequent questions.

[22] We reject the hypothesis that log-income is negative with a 10% level significance.

[23] Not controlling for subsequent questions does not affect our results qualitatively, except for the coefficient on log-income. The coefficient becomes significantly different from zero, but at a 10% level of confidence, and we reject the hypothesis that log-income is negative with a 5% level of significance.

Our evidence suggest that, despite having coherent posterior estimates (and also coherent standard estimates from models (1) and (2)), the underlying heterogeneity reveals a different picture. There are respondents that are not conveying their preferences. We consider only Class 2 satisfies our validity criteria and can be interpreted as providing VSL estimates at face value. It is not surprising to find noisy answers in a self-administered internet survey. What is novel is that we are able to determine the fraction of the our sample that provides nonsensical answers. We find that up to 60% of our sample, can be categorized as providing responses that are not consistent with informed, rational preferences.

5.2.1 Class membership

Table (7) reports on the marginal effects of demographics on class-membership probability.

Table 7: Marginal effects of demographics on Class-membership

	Class 1	Class 2	Class 3
Household with only child	0.04	-0.07***	0.03
	(0.03)	(0.02)	(0.04)
Household with only another Adult	0.05*	-0.04	-0.01
	(0.03)	(0.03)	(0.03)
Household with child + another Adult	0.03	-0.15***	0.12**
	(0.05)	(0.03)	(0.05)
Training succes	0.12**	0	-0.12**
	(0.05)	(0.01)	(0.05)
Log-time	0.03***	0.02**	-0.05***
	(0.01)	(0.01)	(0.01)
Education (High School)	0.10***	0.03	-0.13***
	(0.04)	(0.04)	(0.05)
Education (College)	0.13***	-0.02	-0.11***
	(0.04)	(0.04)	(0.04)
Log-Income	0.02	0.07***	-0.09***
	(0.02)	(0.02)	(0.02)
Environmental concern	-0.02	-0.16***	0.18***
	(0.05)	(0.04)	(0.07)

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. Training success corresponds to not have committed any mistakes during the training sessions. Environmental concern corresponds to a dichotomous self assessed level of importance given to environmental matters where one is equal to high. Robust standard errors in parenthesis.
* significant at 10%; ** significant at 5%; *** significant at 1%

As compared to single person households, households with children or with children and other adults are statistically less likely to belong to Class 2. The latter type of households are consistently more prevalent in Class 3. Households with only another adult are more likely to be found in Class 1.

Success during the probabilistic training phase has no apparent effect on the probability of belonging to Class 2. Rather, success increases the probability of belonging to Class 1 at the expense of reducing the membership probability to Class 3. Time spent completing the survey has a positive impact on Class 2 membership probability, as well as for Class 1.

Regarding education we find that having a high school or college degree has no impact on membership to class 2, but rather increases the odds of belonging to Class 1. The more educated the respondent is, the lower the probability of belonging to Class 3. Moreover, income has a positive impact on the membership to the Class 2, while it has a negative impact on the membership to Class 3. Environmental concern has a negative impact on the membership of Class 2 and increases membership to Class 3.

5.2.2 Willingness to pay

To fix ideas, a graphical representation of the LCA in the log-normal scale is provided in Figure (1). The fine line corresponds to the estimated standard log-willingness to pay, while the bold line corresponds to the estimated latent class log-willingness to pay. The difference between, Class 1, Class 2 and Class 3 is apparent. Class 2 is on the far left, while Class 3 is on the far right of Figure (1).

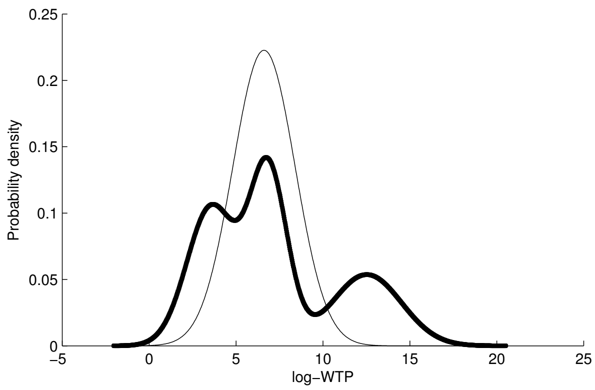

Figure 1: Log-WTP, Gaussian Mixture versus a standard normal assumption

Table (8) reports on the mean value per statistical life based on Class 2 estimates alone. We consider Class 2 as the only sub-group of respondents fulfilling theoretical expectations.[24] We report VSL, as well as SR-VSL, which are virtually the same. Point estimates mean VSL for children are between 6.11 and 6.67 million €, while estimates

[24]Note that we are taking the mean estimates and not the median estimates as in Table (5). This comes from the fact that most of the heterogeneity previously captured by the variance in the standard model is controlled for when performing the LCA. Figure (1) clearly shows the reason why the variance is so large in the standard model; as the heavy line, representing LCA, is found to have a better fit to the data. In fact, the mean VSL under the standard model easily surpasses 2000 million €, a feature not unique to this survey, but rather a common feature in CV. A similar, unreasonably high, mean VSL can be found in Hammitt & Haninger (2010). In the latter paper, only the median is reported. Hanhemann et al. (1991) advise in favour of median VSL, given that it is robust to outliers. The average mean to median ratio found in Hammitt & Haninger (2010) is 350.

for another adult within the household are between 5.56 and 6.15 million €. Finally, mean VSL for adults, from risk addressed to themselves is around 2.24 to 3.39 million €. The largest difference between children and adult mean VSL is on the order of 3. It is interesting to note that median SR-VSL for children, another adult and self (in table 5) are close to the mean VSL computed from Class 2.

Table 8: Mean value per statistical life: Class 2

	VSL	SR-VSL	Mean/Median
VSL Self	2.44	2.24	9.24
	(0.59)	(0.94)	
VSL Child	6.66	6.11	9.24
	(2.59)	(3.22)	
VSL Other adult	6.15	5.64	9.24
	(1.66)	(2.47)	
VSL per Household member	0.61	0.56	9.24
	(0.18)	(0.26)	
VSL Self (Single Person Household)	3.39	3.11	9.24
	(1.04)	(1.44)	

Notes: Values are in millions of euros. WTP is calculated using the specification from each model. VSL is estimated for the mean individual in the following way: first, we take the exponential WTP for the mean individual + the variance over two; second, we multiply the predicted WTP by the low risk reduction (1/10,000); For SR-VSL, we multiply by the elasticity of the risk reduction with respect to the WTP for each model. Standard errors are in parenthesis (delta method). The mean VSL is computed by adding variance over two before taking the exponential. Class' 2 mean/median ratio is reported.

Finally, the mean to median ratio is of the order of 9, which is 54 times smaller than the mean over median ratio from the standard model reported in table (5). Moreover, the median from class 2, reported in table (5), is considerably smaller than the median from the standard model in all the cases. The difference is explained by the fact that the standard model is not disentangling individuals from class 1 and class 3, which have high WTP but also do not satisfy our validity criteria.

6 Discussion

Heterogeneity abounds in our survey as illustrated by Figure (2). Each sub-figure represents the kernel density function of individual estimated posterior coefficients, and each distribution is far from being single peaked with a small variance. Clearly, such heterogeneity, if ignored, may lead to invalid conclusions. What can we learn from it and, what should be done to characterize it?

As we have seen in section 5.2, Class 3 individuals, while exhibiting scope sensitivity, income sensitivity, and baseline risk insensitivity have infeasible WTP estimates. A possible explanation can be found in Kahneman et al. (1993), where they allude to a contribution model. The latter paper suggest that "the responses are better described as expressions of attitudes than as indications of economic value, contrary to the assumptions of the contingent valuation method." Moreover, as observed in Table (7), higher environmental concern expressed by the respondents, make it more likely for them to belong to Class 3. It follows that high environmental concern might lead to over-reactions (Patt & Zeckhauser 2000), though in the form of high WTP estimates, and not scope insensitivity as suggested by Sunstein & Zeckhauser (2010).

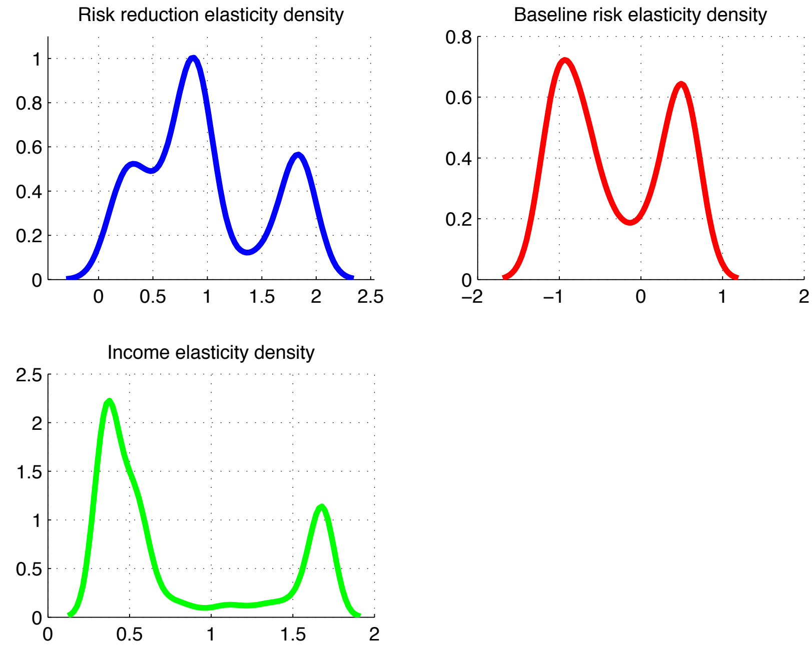

Figure 2: Scope, baseline risk and income elasticity estimated densities

As Class 2 is behaving as predicted by expected-utility theory, should we take into account only their preferences? To assume that these respondents understand the good being valued, is tempting. Sunstein (2013) argues that regulators should use preferences that are informed and rational. From a welfare point of view, Adler (2011) also argues that preferences, which are fully informed and fully rational, should be the ones taken into account. Nevertheless, is Class 2 a representative sample of the French population? Table (9) reports on class-dependent mean demographics. There are no large differences between Class 2 respondents and the full sample. The average age for respondents in Class 2 is 42, half of them are women, half have at least a high school degree, and earn on average 2865 € per month. So, if we believe that Class 2 is sensible enough, we must then choose the estimate for VSL accordingly: 6.11 and 6.67 million € for children, and 2.24 to 3.39 million € for adults.

As seen previously, Lui et al. (2006) shows that allowing for self-protection might lead WTP to decrease with higher levels of baseline risk. Under the Lui et al. (2006) setting, the relationship between risk reduction and WTP is that of near-proportionality, for small risk reductions. Under such circumstances, our Class 1 could be admitted as a plausibly exhibiting the true preferences. Table (8) reports mean VSL, and mean SR-VSL from

Table 9: Demographics conditional on Class-membership

Class	Income	Age	Gender	High School	College	N
1	3036.86	42.85	0.51	0.58	0.26	
	(1665.37)	(13.58)	(0.5)	(0.49)	(0.43)	292
2	2865.42	42.00	0.50	0.51	0.27	
	(1657.59)	(12.48)	(0.5)	(0.5)	(0.41)	415
3	2754.96	44.19	0.55	0.46	0.23	
	(1699.72)	(13.24)	(0.5)	(0.5)	(0.44)	293
All Sample						
	2884.79	42.92	0.52	0.52	0.26	
	(1675.85)	(13.15)	(0.5)	(0.5)	(0.43)	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 male = 0, college corresponds to having up to a college degree, and high school (HS) corresponds to having only a high school degree.

Class 1. As compared to Class 2, mean VSL are higher for Class 1 in all cases. When controlling for the lack of scope sensitivity present in Class 1, the SR-VSL point estimates are quite similar as the SR-VSL from Class 2. Nevertheless, the wide standard errors reflect the lack of scope sensitivity.

Sustein (2013) argues 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." In this paper, we propose two ways to deal with such failures: (1) to investigate preference heterogeneity in a way which allows the researcher to disentangle respondents who are revealing their economic preferences;[25] (2) to implement the scope-revealing VSL. The simplicity of SR-VSL is it greatest appeal.

While scope insensitivity appears to be the norm in CV, VSL is computed and interpreted using an economic model that predicts near-proportionality. Standard median VSL produces estimates that are robust to over reacting respondents, like those found in Class 3, yet, it is not robust to lack of scope-sensitivity. Accounting for the lack of scope sensitivity is necessary, and median SR-VSL can be used for that purpose.

The result of this paper is consistent with other literature where VSL computed from WTP for personal risk reductions is lower, than VSL assessed from WTP for risk reduction to others. The literature suggests that differences between own VSL and a VSL for a child can be explained by age (Chanel et al. 2004, Aldy et al. 2008), risk perception (Hammitt et al. 2004), context of valuation (altruism), and different perspective (society, children or parental). Empirical studies suggest that perspective and altruism substantially influence WTP (Dickie & Ulery 2001). While the differences between children and adults might

[25] We use Latent Class Analysis for this purpose, but it is just one of other strategies a researcher could use.

Table 10: Mean value per statistical life Class 1

	VSL	SR-VSL	Mean/Median
VSL Self	17.32	3.16	1.87
	(1.82)	(2.26)	
VSL Child	27.67	5.05	1.87
	(4.73)	(3.71)	
VSL Other adult	32.90	6.01	1.87
	(3.70)	(4.31)	
VSL per Household member	9.66	1.76	1.87
	(0.85)	(1.25)	
VSL Self (Single Person Household)	26.51	4.84	1.87
	(3.56)	(3.50)	

Notes: Values are in millions of euros. WTP is calculated using the specification from each model. VSL is estimated for the mean individual in the following way: first, we take the exponential WTP for the mean individual + the variance over two; second, we multiply the predicted WTP by the low risk reduction (1/10,000); For SR-VSL, we multiply by the elasticity of the risk reduction with respect to the WTP for each model. Standard errors are in parenthesis (delta method). The mean VSL is computed by adding variance over two before taking the exponential. Only the mean over median ratio is reported.

not appear problematic, the difference between VSL for another adult and for oneself is. Controlling for individual heterogeneity, the difference in VSL maybe explained by altruism. Which should we take? The own VSL, or the altruism augmented VSL? [26]

While in other areas of economics introducing heterogeneity is key in solving issues,[27] not much attention has been given to it in the CVM literature. By introducing heterogeneity in the analysis, our results suggest that fewer than half of our sample satisfy our theoretical validity checks, while the other half is considered as not revealing their preferences. We base our results on the, still representative, sub-group of respondents from which we properly elicit their preferences. Finally, we introduce a novel way to correct for the quality of respondents' answers when computing the value per statistical life, the scope-revealing value per statistical life.

[26] Bergstrom (2004) states that VSL should be estimated over own risk reductions since respondents are better informed over their own preferences.

[27] For example, in Industrial Organization introducing heterogeneity is essential when analysing consumer demands, since it allows to break the Independence of Irrelevant Assumption (or IIA) implicitly introduced by the Logit setting.

References

[1] Adler, M. *Well-being and fair distribution: beyond cost-benefit analysis.* Oxford University Press, 2012.

[2] Alberini, A. Efficiency vs bias of willingness-to-pay estimates: bivariate and interval-data models. *Journal of Environmental Economics and Management 29*, 2 (1995), 169–180.

[3] Alberini, A., Hunt, A., and Markandya, A. Willingness to pay to reduce mortality risks: evidence from a three-country contingent valuation study. *Environmental and Resource Economics 33*, 2 (2006), 251–264.

[4] Alberini, A., Krupnick, A., Cropper, M., Simon, N. B., and Cook, J. The willingness to pay for mortality risk reductions: A comparison of the united states and canada. *FEEM Discussion Paper 92.01, Milan, Italy* (2001).

[5] Aldy, J. E., and Viscusi, W. K. Adjusting the value of a statistical life for age and cohort effects. *The Review of Economics and Statistics 90*, 3 (2008), 573–581.

[6] Arrow, K., Solow, R., et al. Report of the NOAA panel on contingent valuation, 1993.

[7] Arrow, K. J. Essays in the theory of risk-bearing. *The Journal of Business 47*, 1 (1974), 96–98.

[8] Banerjee, S., and Murphy, J. H. The scope test revisited. *Applied Economics Letters 12*, 10 (2005), 613–617.

[9] Breyer, F., and Felder, S. The dead-anyway effect revis (it) ed. Tech. rep., CESifo Working Paper, 2002.

[10] Cameron, T. A., and DeShazo, J. Differential attention to attributes in utility-theoretic choice models. *Journal of Choice Modelling 3*, 3 (2010), 73–115.

[11] Campbell, J. Y. *Consumption-based asset pricing*, vol. 1. Handbook of the Economics of Finance, 2003.

[12] Carson, R. T. Contingent valuation: a practical alternative when prices aren't available. *The Journal of Economic Perspectives 26*, 4 (2012), 27–42.

[13] Carson, R. T., and Mitchell, R. C. *Using surveys to value public goods: The contingent valuation method.* Resources For the Future, 1989.

[14] Carson, T. R., and Groves, T. Incentive and informational properties of preference questions. *Environmental Ressource Economics 37*, 1 (2007), 181–210.

[15] Chanel, O., Luchini, S., et al. Monetary values for air pollution risk of death: A contingent valuation survey. <*halshs-00272776*> (2008).

[16] Chetty, R. A new method of estimating risk aversion. *The American Economic Review 96*, 5 (2006), 1821–1834.

[17] Corso, P. S., Hammitt, J. K., and Graham, J. D. Valuing mortality-risk reduction: using visual aids to improve the validity of contingent valuation. *Journal of Risk and Uncertainty 23*, 2 (2001), 165–184.

[18] Dempster, A. P., Laird, N. M., and Rubin, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological) 39*, 1 (1977), 1–38.

[19] Desaigues, B., Rabl, A., Ami, D., My Kene, B., Masson, S., Salomon, M.-A., and Santoni, L. 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 (2007), 675–698.

[20] Diamond, P. A., and Hausman, J. A. Contingent valuation: Is some number better than no number? *The Journal of Economic Perspectives 8*, 4 (1994), 45–64.

[21] Dickie, M., and Ulery, V. L. 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, 2001.

[22] Doucouliagos, H., Stanley, T., and Viscusi, W. K. Publication selection and the income elasticity of the value of a statistical life. *Journal of Health Economics 33*, 1 (2014), 67–75.

[23] Eeckhoudt, L. R., and Hammitt, J. K. Background risks and the value of a statistical life. *Journal of Risk and Uncertainty 23*, 3 (2001), 261–279.

[24] Evans, M. F., and Smith, V. K. Measuring how risk tradeoffs adjust with income. *Journal of Risk and Uncertainty 40*, 1 (2010), 33–55.

[25] Hammitt, J. K. Valuing mortality risk: theory and practice. *Environmental Science & Technology 34*, 8 (2000), 1396–1400.

[26] Hammitt, J. K., and Graham, J. D. Willingness to pay for health protection: inadequate sensitivity to probability? *Journal of Risk and Uncertainty 18*, 1 (1999), 33–62.

[27] Hammitt, J. K., and Haninger, K. Valuing fatal risks to children and adults: Effects of disease, latency, and risk aversion. *Journal of Risk and Uncertainty 40*, 1 (2010), 57–83.

[28] Hammitt, J. K., and Liu, J.-T. Effects of disease type and latency on the value of mortality risk. *Journal of Risk and Uncertainty 28*, 1 (2004), 73–95.

[29] Hammitt, J. K., and Robinson, L. A. The income elasticity of the value per statistical life: Transferring estimates between high and low income populations. *Journal of Benefit-Cost Analysis 2*, 1 (2015), 1–29.

[30] Hanemann, M., Loomis, J., and Kanninen, B. Statistical efficiency of double-bounded dichotomous choice contingent valuation. *American Journal of Agricultural Economics 73*, 4 (1991), 1255–1263.

[31] Hanemann, W. M. Valuing the environment through contingent valuation. *The Journal of Economic Perspectives 8*, 4 (1994), 19–43.

[32] Hausman, J. Contingent valuation: from dubious to hopeless. *The Journal of Economic Perspectives 26*, 4 (2012), 43–56.

[33] Hess, S., Ben-Akiva, M., Gopinath, D., and Walker, J. Advantages of latent class over continuous mixture of logit models. *Institute for Transport Studies, University of Leeds. Working paper* (2011).

[34] Jones-Lee, M. The value of changes in the probability of death or injury. *The Journal of Political Economy 82*, 4 (1974), 835–849.

[35] Kahneman, D. *Thinking, fast and slow.* Macmillan, 2011.

[36] Kahneman, D., and Knetsch, J. L. Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management 22*, 1 (1992), 57–70.

[37] Kahneman, D., and Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society 47*, 2 (1979), 263–291.

[38] Kaplow, L. The value of a statistical life and the coefficient of relative risk aversion. *Journal of Risk and Uncertainty 31*, 1 (2005), 23–34.

[39] Liu, L., and Neilson, W. S. Endogenous private safety investment and the willingness to pay for mortality risk reductions. *European Economic Review 50*, 8 (2006), 2063–2074.

[40] Machina, M. J. *Choice Under Uncertainty: Problems Solved and Unsolved*, vol. 14. Springer Netherlands, 1992.

[41] Nielsen, J. S., and Kjær, T. Does question order influence sensitivity to scope? empirical findings from a web-based contingent valuation study. *Journal of Environmental Planning and Management 54*, 3 (2011), 369–381.

[42] Oken, E., Choi, A. L., Karagas, M. R., Mariën, K., Rheinberger, C. M., Schoeny, R., Sunderland, E., and Korrick, S. Which fish should I eat? Perspectives influencing fish consumption choices. *Environmental Health Perspectives 120*, 6 (2012), 790.

[43] Patt, A., and Zeckhauser, R. Action bias and environmental decisions. *Journal of Risk and Uncertainty 21*, 1 (2000), 45–72.

[44] Pratt, J. W., and Zeckhauser, R. J. Willingness to pay and the distribution of risk and wealth. *Journal of Political Economy 104*, 4 (1996), 747–763.

[45] Rubinstein, A. Response time and decision making: an experimental study. *Judgment and Decision Making 8*, 5 (2013), 540–551.

[46] Smith, V. K., and Desvousges, W. H. An empirical analysis of the economic value of risk changes. *The Journal of Political Economy 95*, 4 (1987), 89–114.

[47] Sunstein, C. R. The value of a statistical life: some clarifications and puzzles. *Journal of Benefit-Cost Analysis 4*, 2 (2013), 237–261.

[48] Sunstein, C. R., and Zeckhauser, R. Dreadful possibilities, neglected probabilities. *The Irrational Economist, Making Decisions in a Dangerous World* (2010).

[49] Train, K. E. *Discrete choice methods with simulation.* Cambridge University Press, 2009.

[50] Viscusi, W. K. Prospective reference theory: Toward an explanation of the paradoxes. *Journal of Risk and Uncertainty 2*, 3 (1989), 235–263.

[51] Viscusi, W. K., and Aldy, J. E. The value of a statistical life: a critical review of market estimates throughout the world. *Journal of Risk and Uncertainty 27*, 1 (2003), 5–76.

[52] Watt, R., and Vázquez, F. J. The effect of prices on risk aversion. *Theoretical Economics Letters 2*, 1 (2012), 40.

Appendix

Consider an individual who faces two states of the world: either he lives, or he dies. The living state occurs with probability π, and the dying state with the complementarity probability. If he lives, the individual will enjoy wealth, w, and if he dies we will assume that he will be able (willing) to bequeath his wealth to his dependents. The individual is assumed to behave as an "expected utility maximizer" (Jones-Lee 1974), where he selects the decision which maximizes his expected utility given by:

$$E\mathcal{U} = \pi u_A(w) + (1-\pi) u_D(w)$$

where $u_j(w)$ is the utility associated with wealth w conditional on the state of the world j, where $j = \{A, D\}$ corresponds to alive and dead, respectively. In both states, individuals will be assumed to prefer more wealth to less, $\frac{\partial u_j(w)}{\partial w} > 0$, and to be financially risk averse, $\frac{\partial^2 u_j(w)}{\partial^2 w} \leq 0$, for $j = \{A, D\}$.

Consider now that the individual is offered an opportunity to increase his probability of survival, π, by an amount e. In turn for the increase of his survival probability, the individual is willing to forfeit an amount $P(e, w, \pi)$. This amount is, by definition, one that leaves the individual with the same expected utility as with the initial survival probability. The amount, $P(e, w, \pi)$, is defined as:

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

Here, $P(e, w, \pi)$ is the compensating variation in wealth for a change in probability e, (Jones-Lee 1974). In what follows, we use the following notation:

$$
\begin{aligned}
P(e,w,\pi) &\equiv P\\
P(0,w,\pi) &\equiv P_0\\
u_j(w-P(e,w,\pi)) &\equiv u_j(w_e)\\
\frac{\partial u_j(\cdot)}{\partial w} &\equiv u_j'(\cdot)\\
\frac{\partial^2 u_j(\cdot)}{\partial^2 w} &\equiv u_j''(\cdot)\\
(\pi+e)\,u_A^{'}(w_e)+(1-\pi-e)\,u_D^{'}(w_e) &\equiv E\mathcal{U}'(w_e)\\
(\pi+e)\,u_A^{''}(w_e)+(1-\pi-e)\,u_D^{''}(w_e) &\equiv E\mathcal{U}''(w_e)\\
\pi u_A^{'}(w)+(1-\pi)\,u_D^{'}(w) &\equiv E\mathcal{U}'(w)\\
\pi u_A^{''}(w)+(1-\pi)\,u_D^{''}(w) &\equiv E\mathcal{U}''(w).
\end{aligned}
$$

Elasticity of willingness to pay with respect to risk reduction.

To investigate the relationship between P and e we first differentiate 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 $e\to 0$ we have $\frac{\partial P}{\partial e}\equiv VSL$, and $P=0$. If we multiply by e and divide by P we have:

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

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

$$
\lim_{e\to 0}\eta_e^{wtp}=\lim_{e\to 0}\frac{-e\frac{\partial P}{\partial e}\left(u_A^{'}(w_e)-u_D^{'}(w_e)\right)+u_A(w_e)-u_D(w_e)}{\frac{\partial P}{\partial e}E\mathcal{U}'(w_e)+P\left[u_A^{'}(w_e)-u_D^{'}(w_e)-\frac{\partial P}{\partial e}E\mathcal{U}''(w_e)\right]},
$$

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

$$
\lim_{e\to 0}\eta_e^{wtp}=\frac{u_A(w)-u_D(w_e)}{E\mathcal{U}'(w)\frac{u_A(w)-u_D(w)}{E\mathcal{U}'(w)}}=1.
$$

The relationship between willingness to pay,P, and risk reduction e, when $e\to 0$ 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:

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

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

$$\eta_{1-\pi}^{wtp}=\frac{u_A\left(w_e\right)-u_D\left(w_e\right)-u_A\left(w\right)-u_D\left(w\right)}{\left(\pi+e\right)u_A^{'}\left(w_e\right)+\left(1-\pi-e\right)u_D^{'}\left(w-P\left(e,w,\pi\right)\right)}\frac{1-\pi}{P}.$$

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

$$\lim_{e\to 0}\eta_{1-\pi}^{wtp}=\lim_{e\to 0}\frac{-\frac{\partial P}{\partial e}\left(u_A^{'}\left(w_e\right)-u_D^{'}\left(w_e\right)\right)\left(1-\pi\right)}{\frac{\partial P}{\partial e}E\mathcal{U}'(w_e)+P\left[u_A^{'}\left(w_e\right)-u_D^{'}\left(w_e\right)-\frac{\partial P}{\partial e}E\mathcal{U}''(w_e)\right]}.$$

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

$$\frac{1-\pi}{\pi}\geq\lim_{e\to 0}\eta_{1-\pi}^{wtp}=\frac{1-\pi}{\pi+\frac{u_A^{'}(w)}{u_A^{'}(w)-u_D^{'}(w)}-1}>0.$$

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.

Finally, we investigate the relationship between income and willingness to pay. Differentiating P with respect to w yields

$$\frac{\partial P}{\partial w}=1-\frac{\pi u_A^{'}\left(w\right)+\left(1-\pi\right)u_D^{'}\left(w\right)}{\left(\pi+e\right)u_A^{'}\left(w_e\right)+\left(1-\pi-e\right)u_D^{'}\left(w_e\right).}$$

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

$$\eta_w^{wtp}=\frac{E\mathcal{U}'(w_e)-E\mathcal{U}'(w)}{E\mathcal{U}'(w_e)}\frac{w}{P}.$$

Here, η_w^{wtp} denotes the elasticity of substitution between income and willingness to pay. As we are interested in cases when $e \to 0$ applying l'Hôpital's rule yields:

$$\lim_{e\to 0} \eta_w^{wtp} = \lim_{e\to 0} w \frac{u'_A(w_e) - u'_D(w_e) - E\mathcal{U}''(w_e)\frac{\partial P}{\partial e}}{\frac{\partial P}{\partial e}\mathcal{U}'(w_e) + P\left[\frac{\partial P}{\partial e}E\mathcal{U}''(w_e) + u'_A(w_e) - u'_D(w_e)\right]}.$$

which in turns yields,

$$\lim_{e\to 0} \eta_w^{wtp} = w\frac{u'_A - u'_D}{u_A - u_D} - w\frac{E\mathcal{U}''(w)}{E\mathcal{U}'(w)} > 0$$

Here, η_w^{VSL} corresponds to the elasticity of substitution between the value per statistical life (VSL) and income.

Rationale for the scope-revealing VSL

Consider an individual with baseline survival probability, π, and income, w, such that, conditional on surviving, her utility is $u(w)$. Now, suppose she is offered a risk reduction of size, e_1. We know that she is willing to pay an amount, $P(e_1, w, \pi)$. Suppose, she is also offered a risk reduction of size e_2, for which she is willing to pay an amount $P(e_2, w, \pi)$. By equation (1), if $e_2 > e_1$ then it must be the case that $P(e_2, w, \pi) > P(e_1, w, \pi)$. Then, by the mean value theorem, there exists a risk reduction e_c, such that:

$$\frac{\partial P(e_c, w, \pi)}{\partial e} = \frac{P(e_2, w, \pi) - P(e_1, w, \pi)}{e_2 - e_1}. \tag{16}$$

where the risk reduction, e_c, is bounded between $[e_1, e_2]$. Let η denote the ratio between a percentage change in WTP, and a percentage change in risk. By equation (16), we have:

$$\eta = \frac{\partial P(e_c, w, \pi)}{\partial e}\frac{e_c}{P(e_c, w, \pi)} \approx \frac{\partial P(e_c, w, \pi)}{\partial e}\frac{e_1}{P(e_1, w, \pi)}$$

and rearranging it we find that:

$$\frac{\partial P(e_c, w, \pi)}{\partial e} \approx \eta\frac{P(e_1, w, \pi)}{e_1}. \tag{17}$$

As all the terms in the RHS of equation (17) are empirically available, we are able to compute the marginal rate of substitution between a risk reduction, e_c, and wealth w.

Figure (3) illustrates the above reasoning. The fine line represents, $VSL \equiv \frac{\partial P(0,w,\pi)}{\partial e}$, as defined by the standard model. There are, at least, three ways of approximating VSL:

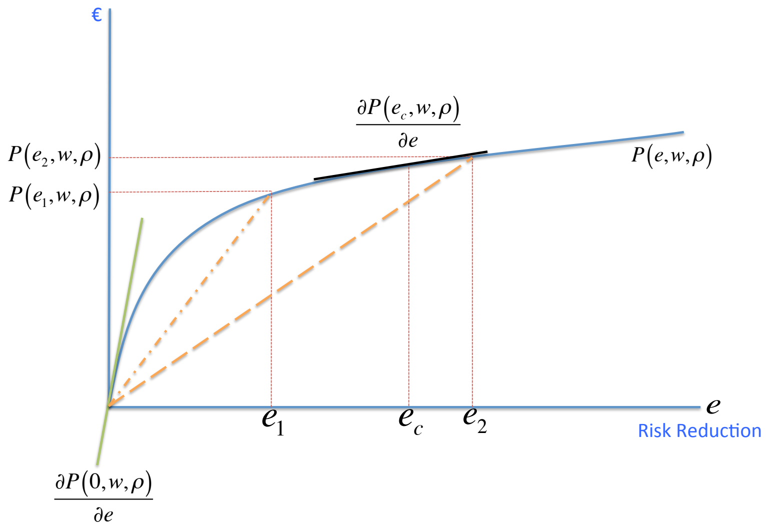

Figure 3: Willingness to pay and scope-revealing value per statistical life

$$VSL \approx VSL_{e_1} \equiv \frac{P(e_1, w, \pi)}{e_1}$$
$$VSL \approx VSL_{e_2} \equiv \frac{P(e_2, w, \pi)}{e_2}$$

where VSL_{e_j} corresponds to the VSL computed from the risk reduction $j = \{1, 2\}$. Their empirical simplicity is their virtue. An alternative way of approximating VSL is given by:

$$VSL \approx VSL_{e_c} \equiv \eta VSL_1$$

which corresponds to VSL_{e_1} corrected by the estimated elasticity, η.

As $e_2 \to 0$ and $e_2 > e_1$, then $\frac{P(e_1,w,\pi)}{e_1}$ in equation (17) should tend to VSL.[28]. When the risks are not zero, the ratio $\frac{P(e_1,w,\pi)}{e_1}$ approximates VSL. In Figure (3), this corresponds to the left-most yellow dotted chord. When $e_1 \to 0$, the yellow dotted chord approximates the green tangent, VSL. Moreover, $e_2 \to 0$ and $e_2 > e_c > e_1$, implies that $\frac{P(e_1,w,\pi)}{e_1} \approx \frac{P(e_2,w,\pi)}{e_2} \approx \frac{\partial P(e_c,w,\pi)}{\partial e}$. But, given the concavity of u, hence of P, as long as, $e_1 > 0$ we have:

[28]The complications occur in communication, and understanding of such small risks. For this reason, the literature has develop strategies to prevent communication, and understanding issues (Corso e al. 2004). A parallel can be established when estimating discount factors. If there is a strong present bias, it might be better to propose to an individual to choose between the future, and a further future.

$$\frac{\partial P\left(0, w, \pi\right)}{\partial e} > \frac{\partial P\left(e_c, w, \pi\right)}{\partial e}. \tag{18}$$

Nevertheless, as $e_2 \to 0$, we have:

$$\frac{\partial P\left(e_c, w, \pi\right)}{\partial e} \to \frac{\partial P\left(0, w, \pi\right)}{\partial e}. \tag{19}$$

Equation (18) shows the scope-revealing VSL is a lower bound of VSL, and from equation (19) we know that the difference tends to zero.

Under an expected utility framework, the three measures are, for a small enough risk reduction, approximately the same. The advantage of the scope-revealing VSL over the other measures is that it accounts for scope-sensitivity in respondents' answers. The correction occurs regardless of the reasons behind the lack of scope sensitivity; poor understanding of the hypothetical good, high risk aversion group of individuals, or others.

Finally, let $w_{e_c} = w - P\left(e_c, w, \pi\right)$. Then we have:

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

It can be re-expressed as:

$$\frac{\partial P\left(e_c, w, \pi\right)}{\partial e} = \frac{\partial P\left(0, w_{e_c}, \pi + e_c\right)}{\partial e} \equiv VSL_{e_c} \tag{20}$$

where VSL_{e_c} denotes the value per statistical life at wealth w_{e_c} and baseline risk $\pi + e_c$. It follows from (20) that VSL_{e_c} is an equally valid measure of the value per statistical life.

Additional regression tables

The reference class is taken to be Class 2, so that all coefficients are to be understood as deviations from class 2 coefficients. The table reports on the influence of environmental concern, time spent completing the survey, whether respondents answered the training questions without error and respondents' income on class-membership.

Table 11: Membership demographics

	Class 1	**Class 3**
Household with only child	0.326	0.283
(Yes=1, 0 o.w.)	(2.74)**	(1.83)*
Household with only another Adult	0.252	0.086
(Yes=1, 0 o.w.)	(1.94)*	(0.52)
Household with child + another Adult	0.503	0.776
(Yes=1, 0 o.w.)	(4.57)**	(5.56)**
Log-time spent on survey	0.069	-0.216
(minutes)	(2.81)**	(5.20)**
Training success	0.455	-0.363
(Yes=1, 0 o.w.)	(4.57)**	(3.51)**
Education	0.236	-0.557
(High School=1, 0 o.w.)	(2.12)*	(4.63)**
Education	0.478	-0.302
(College=1, 0 o.w.)	(4.45)**	(2.61)**
Log-Income	-0.103	-0.496
(log-€)	(1.27)	(4.75)**
Environmental concern	0.33	1.093
(High=1, 0 o.w.)	(4.66)**	(12.63)**
Constant	-1.666	4.958
	(2.57)*	(5.58)**

Notes: Dependent variable are the prior probabilities obtained at each iteration threw the EM algorithm. The estimates are obtained using a fractional logit Identifying assumptions require Class 2 coefficients to be normalized to zero, so the coefficients are to be understood as deviations from Class 2 coefficients. Robust t-statistics in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%