UNIVERSITY OF EXETER | LAND, ENVIRONMENT, ECONOMICS AND POLICY INSTITUTE

Working paper: *07/2018*

The Value of Statistical Life for Adults and Children: Empirical and Methodological Investigations

Ben Balmford, Ian J. Bateman, Katherine Bolt, Brett Day and Silvia Ferrini

Keywords:
Value of statistical life; Contingent valuation; Chaining method; Non-market valuation; Health.

JEL codes:
D60, I10, H40, H51

The Value of Statistical Life for Adults and Children:

Empirical and Methodological Investigations

Ben Balmford[1], Ian J. Bateman[1,2], Katherine Bolt[3], Brett Day[1] and Silvia Ferrini[4]

Abstract

Estimates of the Value of Statistical Life (VSL) provide a vital input to a variety of policy decisions ranging from health provision to transportation planning. However, the bulk of VSL research has focussed on estimating average values rather than taking account of the potential variation in VSL across groups. Policymakers are particularly concerned that using estimates based on data concerning adults might provide poor proxies of the values associated with preventing child fatalities. We investigate this empirical problem while also addressing methodological critiques of standard contingent valuation (CV) approaches to VSL estimation which asks survey respondents to value an outcome described in terms of both the probability of occurrence and the health impact of an event. A prior lab experiment confirms fundamental problems in subjects' abilities to provide internally consistent valuations of such compound goods. Given this we compare CV approaches with the 'chaining method' of Carthy et al. (1999) which splits the valuation task in two, assessing the probability of an event and the disutility of that event separately and then 'chaining' responses together to obtain a VSL estimate. Results confirm prior expectations that VSL values for preventing child fatalities significantly exceed those for adults. However, while we identify many advantages of chaining over CV approaches, through a novel variant of a validation test suggested by Carthy et al. we reveal anomalies in the estimates produced by the chaining method suggesting that a robust method for VSL calculation is yet to be refined.

[1] Land, Environment, Economics and Policy Institute (LEEP), University of Exeter Business School (UEBS), UK.
[2] Corresponding author, email: i.bateman@exeter.ac.uk
[3] Royal Society for the Protection of Birds (RSPB), Sandy, UK.
[4] LEEP and the University of East Anglia, UK

Key words:

Value of statistical life; Contingent valuation; Chaining method; Non-market valuation; Health.

Highlights:

Reports empirical estimates of the Value of Statistical Life for adults and children;

Compares the contingent valuation and chaining methods;

Uses a mix of lab experiment and field survey approaches;

Tests for and finds evidence of anomalies across results provided by both methods.

Introduction

The standard procedure for assessing the economic case for or against public sector interventions is to undertake a cost-benefit analysis, weighing the different costs and benefits in monetary terms. When such interventions involve changes in the risks of death, illness or injury, it is then necessary to find some way of placing a monetary value on these changes. Two rather different approaches have been used. In policy relating to safety, where the focus has been upon accidents which may cause injury and/or premature death, a number of governments and their agencies have used the Value of Statistical Life (VSL) to represent the aggregate of many individuals' willingness to pay (WTP) to avoid small mortality risks that are then used to value the prevention of one expected (in the statistical sense) fatality[5]. In policy relating to health care interventions, where the focus is upon preventing or treating illness that may lead to reductions in health status and/or premature death, an alternative approach has involved measuring the benefits in terms of the Quality Adjusted Life Years (QALYs) gained as a result of an intervention, and then deciding whether the 'cost-per-QALY' is above or below some threshold which is regarded as good value for money[6]. Unlike the VSL approach, using QALYs does not assess welfare benefits, but rather how an intervention compares to some cost threshold.

In contrast to QALYs, in policy decision making the VSL is typically used as if it is unresponsive to age, and may therefore not be a true reflection of preferences (Shepard and Zeckhauser, 1984; Jones-Lee,

[5] An alternative term for essentially the same concept is the Value of Preventing a Fatality (VPF). With either term, the important point to note is that this is *not* the value of preventing the otherwise certain death of an identifiable individual, but the summation of many people's WTP to reduce their own risks by rather small amounts until the total reduction in probabilities adds up to 1. The UK Department for Transport (DfT) values the prevention of a fatality on Britain's roads at £1.83 million (DfT, 2016) although a figure of approximately £1.55m per fatality prevented is listed in its most current guidance for undertaking cost-benefit analyses of road safety schemes https://www.gov.uk/government/publications/webtag-tag-data-book-may-2018. The UK Health and Safety Executive have recently commissioned research into the feasibility and use of the CV based Value of a Life Year (VOLY; Desaigues et al., 2011) concept within UK decision making.

[6] In the UK, for example, the National Institute for Health and Clinical Excellence (NICE) has used per QALY thresholds of less than £20,000 as likely acceptable, and more than £30,000 as in need of good justification in first approximations when judging whether a new health care intervention represents sufficiently good value for money to be adopted by the UK National Health Service. https://www.nice.org.uk/process/pmg6/chapter/assessing-cost-effectiveness and further discussion in Donaldson et al. (2011).

1989; Aldy and Viscusi, 2008). Indeed, preventing the premature death of a child rather than an elderly person will register as a much larger benefit under a QALY-based system which is not reflected in most official VSL measures where the same average value is applied to everyone. This is in part because we may be unable to predict the subset of people whose lives would be saved by accident preventing measures, while the treatment of a particular disease has a clear target population.

Nonetheless, the question of whether the benefit of reducing risk to the elderly should be valued less than corresponding risk reductions for younger groups has become more prominent (see reviews by O'Brien, 2013; and Morgan, 2017). Some countries have contemplated using different VSLs for different age groups, notably: Canada (Hara Associates, 2000), the European Commission (EC, 2001), and, somewhat controversially[7], the U.S. Environmental Protection Agency (Viscusi, 2009). Therefore, it is of interest to know whether people would subscribe to a distinction between the VSLs of children and adults. Although the theoretical and empirical VSL literature is quite extensive (e.g. Alberini, 2005; Hammitt and Zhou 2006), and despite some evidence that age does appear to impact upon the value of preventing a fatality (Aldy and Viscusi, 2008), only a relatively small number of studies specifically address the issue of valuing mortality risks for children (Agee and Crocker 1996; Alberini and Ščasný, 2011; Blomquist et al., 2011; Cropper et al., 2011; Dickie and Gerking 2003; Guerriero et al., 2017; Hammitt and Haninger, 2010; Jenkins et al. 2001; Mount et al 2003; Nastis and Crocker 2003).

Moreover, there is no simple observable monetary value for the VSL (McDaid et al., 2015). Using wages as an estimate of a VSL wrongly equates prices and values (Rice, 2015). Revealed preference techniques using either wage premiums or expenditure on safety equipment (Bellavance et al., 2009) require strong assumptions regarding the information hold on the risk associated with particular jobs

[7] See, for example, the controversy surrounding the US EPA's use of an age-weighted VSL played out in newspaper headlines such as "EPA Drops Age-Based Cost Studies" (New York Times, May 8, 2003), "EPA to Stop 'Death Discount' to Value New Regulations" (Wall Street Journal, May 8, 2003), and "Under Fire, EPA Drops the 'Senior Death Discount'" (Washington Post, May 13, 2003).

or behaviours (Dolan et al., 2008), and values are very sensitive to the exact nature of risk estimation (Scotton, 2013).

These problems have meant that economists frequently apply stated preference (SP) methods such as discrete choice experiments or, most commonly, contingent valuation (CV), to estimate VSLs (e.g. Alolayan et al., 2017; Dickie and Gerking, 2003; Roldós et al., 2017). While CV methods have been used extensively worldwide to estimate willingness to pay (WTP) and other measures for a wide variety of goods (Carson, 2011; Rakotonarivo et al., 2016), the large majority of these applications have been for non-risky options, i.e. goods which, in a contingent market, are certain to be supplied if sufficient funds are paid. As budget constraints, plausibility and ethical principles all mitigate against asking an individual to state their WTP to prevent the certainty of their death from a given cause, this approach is inappropriate for calculating a VSL. Therefore CV studies of health typically value risky options; goods which are provided as probabilities such as a change in a non-unity risk of death or the probability of being afflicted by a disease. These 'compound' goods present CV survey respondents with a difficult challenge: having to simultaneously evaluate (in monetary terms) their value for avoiding some (often unfamiliar) adverse health outcome; and understand the (typically small) probability of that outcome occurring. Both are demanding and unfamiliar tasks and their joint estimation is, arguably, cognitively overwhelming. All CV studies assume that, when stating the value of reducing her[8] mortality risk by a specified amount, an informed individual has well-formed and theoretically consistent preferences (Carson and Groves, 2007). In cases where this assumption does not hold CV responses may be malleable and subject to bias. In particular commentators have long argued that in cases of cognitive overload, respondents may seek to infer information regarding appropriate responses from objectively irrelevant elements of the framing of a valuation question (Nielsen et al., 2012; Slovic, 1995; Tversky and Kahneman, 1973, 1974).

[8] Following convention we adopt the female gender throughout.

Indeed, the SP literature reports a number of persistent anomalies; results which suggest that respondents are unable to relate certain valuation questions to a set of standard economic preferences (Chilton et al., 2004; Desvousges et al., 1992; Dolan et al., 2008; Hausman, 2012; Jones-Lee et al., 1995; Jones-Lee and Loomes 2004; McFadden and Train, 2017). In the context of health outcomes, value estimates have often been found to be insensitive to scope (e.g. inadequately related to changes in the risk of a health state arising; Beattie et al., 1998), even when it cannot be plausibly argued that income is acting as a constraint on WTP (Gyrd-Hansen et al., 2014; Søgaard et al., 2012). Both Jones-Lee et al., (1995) and Dubourg et al., (1997) note inadequate responsiveness in individual's WTP when risks were reduced, resulting in the inflation of corresponding VSL estimates. Clearly insufficient scope sensitivity renders such VSL estimates invalid for decision making purposes as reducing the risk probabilities presented to CV respondents will erroneously drive up the implied VSL. It has been argued that this phenomena is driven in part by the cognitive demands of valuing a compound good (Carthy et al., 1999; Gyrd-Hansen et al., 2012). Given this, we open the applied element of our analysis with a laboratory experiment designed to test how respondents cope with each element of the compound good: 1) valuing outcomes (using both familiar and unfamiliar goods to examine the effects of experience) and 2) assessing small risks of those outcomes occurring (with those risks being varied to examine scope sensitivity).

It was with the particular challenge of compound goods in mind that Carthy et al., (1999) proposed the 'chaining method' to estimate VSLs. This approach splits presentation of the compound good up into a two-step procedure. The first step asks subjects to trade-off a risk of a specified ill-health state against a risk of death (e.g. which is worse, an X% chance of a specified ill-health state or a Y% chance of death; respondents adjust Y until the two outcomes are equivalent). The second step asks the respondent to place a value on avoiding that ill-health state. Combining these responses allows the analyst to 'chain' WTP for the ill-health state up to an inferred WTP to reduce the risk of death and hence the VSL.

The chaining method has been the focus of some debate, critique and defence (e.g. Thomas and Vaughan, 2015; Chilton et al., 2015). The present paper seeks to contribute to this and the wider literature through comparison of the chaining method with the more widely applied CV approach to the estimation of VSL. Furthermore, we use both methods to address the policy relevant question of whether VSLs for children and adults are different. While typically one would only ask someone about risks to their own life, we avoid the cognitive and moral challenges of asking children such demanding questions by investigating the values held by parents for their own lives and those of their children. If such a 'child premium' exists, it should be detectable within the subset of the population who are parents[9]. Any excess of the estimated child VSL over the parent VSL values should provide strong evidence of an age premium associated with young lives.

We also use this analysis to address a methodological challenge. In separating the compound good into its two constituent parts, Carthy et al., argue that the cognitive demand placed on a subject can be much reduced and certainly the results obtained appear promising. We examine this claim by developing a novel variant of a test for internally consistency suggested by Carthy et al. This is achieved by splitting the first stage of the chaining approach into two tasks where respondents initially compare minor with major ill-health state, then compare the latter major ill-health state with death. We term this a 'double' chained method and compare this with the conventional 'single' chain approach pioneered by Carthy et al. Consistency should be evidenced by no significant difference arising in the VSL measures delivered by the single and double chaining variants. Extreme inconsistency will arise if parents apply their 'child premium' at each stage of the chaining exercise, causing resultant VSL estimates to inflate dramatically.

[9] Note, we are not suggesting that any age premium we observe in this group should be used as a social value, but rather that if an age premium does exist in the wider population, then it will be most easily detectable in a sample of parents.

The rest of this paper is organized as follows. First, we discuss and present results from the laboratory experiment testing scope sensitivity of responses across different levels of risk and different levels of good familiarity. Next we present a first field survey of parents comparing the standard Carthy et al. (single chain) approach to the chaining method with a conventional CV analysis of VSL. Finally, we present our consistency test of the chaining method, contrasting the single and double chain variants across a nationwide and nationally representative sample of parents.

Scope sensitivity, familiarity with the good and risk framing: an experiment

As explained above, we wished to test some of the key assumptions inherent in CV studies of VSL in a highly controlled setting. Specifically, the aim of this experiment was to examine the sensitivity of stated WTP responses to: a) the familiarity of the goods being valued; b) the size of the risk reduction offered; and c) the framing of risk probability information.

We conducted this experiment with 99 students at the University of East Anglia, using a self-administered, computerised questionnaire, which randomised the presentation order of treatments and questions[10]. The experimental subjects were presented with three goods of differing levels of familiarity: avoiding losses of money (£75); avoiding a temporary stomach complaint; and avoiding a condition causing temporary blindness. Each of these goods were offered at different levels of risk and probabilities were presented using different formats (either chances in 10, such as 1/10, or changes in 1,000, such as 100/1,000), the latter being a variant of the test for risk framing effects found to be significant by Pinto-Prades et al., (2006). Respondents were asked to value each compound of good, risk and probability presentation in a manner similar to CV VSL studies. The

[10] Further details of the experimental design are available on request from the authors.

question ordering was varied and analysis conducted so as to minimise the potential for initial responses to anchor subsequent responses (Jacowitz and Kahneman, 1995)[11].

Table 1 reports the mean, median and standard deviation values of stated WTP for the various compounds of outcome and risk reduction valued in this experiment. The pattern of values across compound goods appears plausible and panel (a) presents nonparametric tests of the scope sensitivity of WTP within each good. Here the penultimate column presents a series of 'weak' scope sensitivity examining, for each outcome, whether WTP for a substantially (five times) larger risk reduction is significantly greater than that for a smaller risk reduction. In all cases this weak sensitivity test is satisfied. However, the final column of this table tests whether, after scaling up (multiplying by five) WTP for the smaller risk reduction this is insignificantly different to the directly estimated WTP for that larger compound good. In every this 'strong' sensitivity tests finds that the directly estimated WTP is significantly lower than the scaled up WTP for what should be an identical good. This confirms prior results that CV respondents over-estimate WTP for small probability risk reductions (Shogren, 1990; Jones-Lee et al., 1995; Dubourg et al., 1997; Beattie et al., 1998; Chilton et al., 1999; Hammar and Johansson-Stenman, 2004).

[11] It is possible that a subject might use their response to an initial question to calculate their response to a subsequent question rather than referring to their preferences afresh. For example, if a respondent is first asked about her WTP to avoid a 1/10 chance of a stomach bug, she may then simply multiply her response by five to determine her WTP to avoid a 5/10 chance of a stomach bug. The likelihood of such 'anchoring' is potentially reduced where risks are expressed using different denominators (where the necessary calculation is far less obvious, e.g. where the second question concerns a 500/1000 risk), or between different outcomes. Therefore, to avoid these potential problems of anchoring, we focus our analyses on the first responses that an individual gives for a risk expressed using a particular denominator for each negative outcome. In the example above the answer to the 1/10 risk would be included in our analyses, while a subsequent response regarding a 5/10 risk would not because the denominators are identical, whereas a subsequent response to a 500/1000 would be included in analysis. Question ordering was varied so a response to the latter question presented first in the list seen be an individual respondent would also be included in our analysis.

(a) Tests for scope sensitivity	Mean WTP (£) *Median* (St. Dev.)		Weak test for scope sensitivity:* WTP smaller risk v WTP larger risk [Standard theory **does** predict a difference] p value (z statistic)	Strong test for scope sensitivity:* Scaled WTP smaller risk v WTP larger risk [Standard theory **does not** predict a difference] p value (z statistic)
	Smaller risk reduction	Larger risk reduction		
Risk in 10	*1/10 to 0/10*	*5/10 to 0/10*		
Money stolen	6.21 *5.00* (5.69)	19.55 *15.00* (14.36)	<0.001 (-5.56)	0.044 (-1.71)
Stomach bug	8.57 *5.20* (9.05)	23.96 *15.00* (31.9)	<0.001 (-4.39)	0.004 (-2.66)
Temporary blindness	29.23 *10.00* (75.89)	57.27 *30.00* (104.22)	<0.001 (-3.41)	0.026 (-1.95)
Risk in 1,000	*20/1,000 to 0/1,000*	*100/1,000 to 0/1,000*		
Money stolen	5.17 *4.50* (5.19)	14.49 *7.95* (20.28)	<0.001 (-4.04)	0.011 (-2.30)
Stomach bug	7.61 *5.00* (8.02)	14.81 *8.10* (17.03)	0.009 (-2.35)	<0.001 (-3.58)
Temporary blindness	16.98 *6.15* (29.17)	40.58 *19.40* (67.00)	<0.001 (-3.11)	0.033 (-1.84)

(b) Tests for framing effects	Mean WTP (£) *Median* (St. Dev.)		Significance of difference between 0.1 risk framed either as 1/10 or 100/1,000[12]* [Standard theory **does not** predict a difference] p value (z statistic)
Risk framing	*1/10 to 0/10*	*100/1,000 to 0/1,000*	
Money stolen	6.21 *5.00* (5.69)	14.49 *7.95* (20.28)	<0.001 (-3.29)
Stomach bug	8.57 *5.20* (9.05)	14.81 *8.10* (17.03)	0.033 (-1.85)
Temporary blindness	29.23 *10.00* (75.89)	40.58 *19.40* (67.00)	0.041 (-1.74)

Table 1: WTP responses to avoid negative outcomes at different risk levels: Tests for (a) scope sensitivity and (b) framing effects

Note: Heavier weight grid cells denote data rather than test results

* Non-parametric one-tailed Mann-Whitney Wilcoxon test

[12] One tail test examines whether the 1/10 risk is perceived as smaller than the 100/1,000 risk.

Panel (b) of Table 1 tests whether the framing of a 0.1 risk as either 1/10 or 100/1,000 alters WTP. Test results clearly reject equality in all cases with WTP to reduce a 100/1,000 risk consistently and significantly greater than that to reduce a 1/10 risk. This clear evidence of framing effects within such a deliberately straightforward experiment suggest that, when faced with compound, risky options, WTP responses to standard CV questions are likely to fail basic anomaly tests.

These simple tests question the assumptions underpinning the CV approach to valuing compound risky options. The chaining method avoids this challenge by splitting the risk assessment and valuation tasks. In the following section we provide a field study comparison of the standard CV approach to VSL estimation to that provided by the Carthy et al., chaining method.

Comparing the contingent valuation and chaining methods for estimating VSL: A first field study

Our experimental results raise considerable concerns regarding the ability of individuals to provide consistent valuations of compound risky health options. The chaining approach was specifically designed to address such challenge. However, how does it perform relative to the more commonly applied CV method when applied in a relevant, field context and how would both approaches address the policy relevant question of whether VSL varies between adults and children? Our first field study set out to answer these questions.

The chaining method

As summarised previously, the chaining method was first developed by Carthy et al (1999) with the intention of overcoming some of the difficulties faced by the CV method in valuing small changes in health risk. An illustrative example of a CV-style question is given below, with the text in parentheses showing the changes employed when a respondent is asked about their child rather than themselves.

Note that the respondent is being asked to simultaneously consider both the value of avoiding a negative outcome and the risk of that outcome.

> "*Consider a product that you could buy which reduces your (child's) risk of dying over ten years by 5/1000. The product has no other benefits or side effects except reducing the risk of death. Suppose that this product was not provided through public health services, nor would it be covered by private health insurance. Therefore the only way to obtain this product would be for you to pay for it.*
> *What is the maximum amount you would be willing to pay for this product?"*

The chaining method avoids asking respondents to directly value a change in mortality risk, breaking the valuation and risk parts into two steps. The first step essentially uses the CV approach to ask respondents to value the avoidance or cure of what would otherwise be the certainty of a non-fatal ill-health incident. An illustrative example is given below (with the text in parentheses again referring to a respondent being asked about their child rather than themselves).

> "*Imagine that a test shows that you (your child) is going to suffer severe stomach pains, diarrhoea and vomiting for 2-3 days every 2 weeks for 12 months.*
> *I want you to suppose that a treatment is available which would avoid all of the effects of this to you (your child). Suppose that this treatment was not provided through public health services, nor would it be covered by private health insurance. Therefore the only way to obtain this alternative treatment would be for you to pay for it.*
>
> *What is the maximum amount you would be willing to pay for this treatment which would bring you (your child) back to full health within a few days, after which you (your child) would be cured?*"

Adjustments to the payment elicitation format as well as the wording of other parts of the question can be made, but crucially this question involves certain, as opposed to risky, outcomes; the respondent does not have to simultaneously consider their willingness to pay and the probability of an outcome simultaneously.

In the second stage of the chaining method, respondents undertake a risk trade-off. Under the original Standard Gamble approach, a respondent is asked to trade-off between either (i) the certainty of a specified ill-health state or (ii) a treatment which has some chance $(1 - \pi_j)$ of delivering full health and some risk (π_j) of death. Respondents vary π_j until they are indifferent between (i) and (ii). This risk level can then be applied to the respondents WTP to avoid the certainty of the specified ill-health state to obtain their imputed VSL.

While in principle the Standard Gamble approach should provide an unbiased estimate of π_j, Carthy et al., (1999) argue that it may be subject to a "certainty effect" if respondents are unwilling to trade ill health states for anything but negligible mortality risks. At the extreme a complete unwillingness to accept any risk of death, π_j, would lead to an infinite VSL. To alleviate this concern and replicate the Carthy et al approach we employ a Modified Standard Gamble (MSG; Baker and Robinson 2004; Jones-Lee et al 1995). The MSG asks the respondent to imagine she has become unwell and is taken to hospital where doctors tell her that if she is not treated then she is certain to die. However, she is also told that there are two possible treatments available to her, both free of charge, for example as follows:

> *Treatment A: If successful, the treatment will result in the respondent experiencing the consequences of a specified, non-fatal, ill-health state for a defined period* [e.g. the severe stomach pain, diarrhoea and vomiting for 12 months case mentioned

previously, or ill-health states X and Y referred to subsequently]. *However, if the treatment is unsuccessful then the patient would fall unconscious and die shortly afterwards with probability of 1/1,000* [this risk is set by the analyst and can be varied].

Treatment B: If successful, this treatment will result in a return to normal health after a couple of days but if unsuccessful there is a risk that the treatment will result in immediate unconsciousness followed shortly by death [this risk, π_j, is varied until the respondent states that they are indifferent between Treatment A and B].

If successful then Treatment B has a better outcome than Treatment A, and therefore it is expected that the respondent will accept some additional mortality risk for the chance of this better successful outcome. Notice that in the MSG format, both treatments involve some risk of death, the intention being that this will counteract any certainty effect and still avoid the direct valuation of a risky compound good as in the CV approach.

The two steps of the chaining procedure thus provide: a) a link between money and the certainty of a health state; and b) a link between that health state and a risk of death. These are then linked ('chained') together, connecting mortality risk to a money sum, from which a VSL can be derived.

Survey design and sampling

The survey comprised eight main sections: the first four asked the questions necessary to undertake the chaining method and the last four relate to the CV method. The chaining method preceded the direct CV method, as the text describing the chaining method introduced concepts within the survey and included some "warm-up" questions. Randomising the order of the two approaches was neither

possible (as it would have required altering or adding to the survey text, hindering comparability of the results between treatments) nor considered desirable: including CV questions after the chaining method was deemed likely to improve the consistency of responses to the CV questions and therefore, if anything, was likely to favour the standard CV method. Eight different treatments of the survey were used to control for question ordering effects within each of the two methods (see Appendix 1). Respondents were all parents recruited from the Cambridgeshire area, approached either through schools or at local attractions for young children. The sample was not selected to be representative of parents nationally; rather, by randomly allocating each parent to each treatment, we sought to test how robust the CV and chaining methods were to anomalies. In total, 300 respondents took part in this survey.

Treatments used to describe the non-fatal ill-health scenario were introduced in the warm-up section and parents were asked to consider the following ill-health states[13], X and Y as follows:

- X: 3 weeks hospitalisation; 4 months severe pain; permanent pain in hip
- Y: 2 months hospitalisation; 4 months moderate pain; permanent pain in knee

After reading through the ill-health states the respondent ranked them in terms of perceived disutility first if she was affected and then if the ill-health states were suffered by her child.

In the chaining method section, parents were asked separate open-ended questions about their WTP for the certainty of avoiding each of these ill-health states and MSG questions relating each of these states to risks of death. Each respondent was asked to answer on her own behalf and on behalf of one randomly selected child in her household[14]. Thus in the first step of the chaining method we obtained

[13] Pre-testing also examined a third ill health state, Z, described as: "2 weeks hospitalisation; 2 months severe pain and bedridden; no permanent ill-health." However, piloting revealed that respondents were not willing to trade the risk of ill-health and mortality risk for their children in the MSG with the ill-health state Z, and it was therefore decided to focus these questions on various combinations involving just X and Y.

[14] If the respondent had more than one child, she was asked to consider the child whose birthday was next.

WTP values for adult and child to avoid the certainty of the ill-health states described above; and from the MSG we identified the risk of death in Treatment B at which the respondent was indifferent between Treatment A and Treatment B for herself and, separately, also for her child.

The questions for the CV part of the questionnaire were based on the survey design employed in the highly cited Krupnick et al. (2002) study. After some questions intended to help the respondent think about her overall baseline mortality risk, she was asked to consider how much she would value a product that would reduce her risk of dying over the next ten years. Each respondent was asked about products which would reduce the risk of dying over 10 years by 5/1,000 and separately for a risk reduction of 1/1,000; and both questions were also asked with respect to her child. The ordering of the questions were varied depending upon the survey version implemented (see Appendix 1). Each respondent was asked about both levels of risk reduction in order that we could test for possible ordering effects within subject.

Results

Table 2 presents summary statistics of the results for the first step in the chaining method: valuing the certainty of avoiding a negative health outcome. Substantially different mean and median values for a given health outcome and high standard deviations highlight the positive skew of responses. Tests reveal that values for reducing risks to children are very substantially higher than those for adults; a finding which accords with policy concerns.

Table 3 presents summary statistics for the second step in the chained approach, responses to the MSG question, and tests whether these values are significantly lower for children than for adults. Indeed, it is clear that respondents are prepared to accept substantially higher risk levels for themselves than for their children; a finding which accords with the policy concerns motivating this study.

WTP to avoid negative health outcome (£)	Mean (£) *Median* (St. Dev.)		Is the adult value lower than the corresponding child value?* p value (z statistic)
	Adult	Child	
Ill-health state X	29,083 *5,000* (102,317)	97,849 *15,000* (224,334)	<0.001 (-5.94)
Ill-health state Y	16,738 *5,000* (81,824)	112,293 *20,000* (242,255)	<0.001 (-7.11)

Table 2: First step (valuation) of the chaining method: Mean and median WTP values for avoiding the certainty of specified ill-health states

Note: ill-health state X = 3 weeks hospitalisation; 4 months severe pain; permanent pain in hip
ill-health state Y = 2 months hospitalisation; 4 months moderate pain; permanent pain in knee
* Non-parametric one-tailed Mann-Whitney Wilcoxon test

Acceptable π /1,000	Mean *Median* (St. Dev.)		*Is the adult value higher than the corresponding child value?** p value (z statistic)
	Adult	Child	
Ill-health state X	62.47 *25.00* (95.66)	45.12 *10.00* (82.67)	0.006 (-2.50)
Ill-health state Y	75.79 *25.00* (119.99)	42.70 *10.00* (81.30)	<0.001 (-3.17)

Table 3: Second step (MSG) of the chaining method: Levels of mortality risk (π_j) stated by respondent at which they are indifferent between Treatment B and a particular ill-health state (X or Y) which itself has a mortality risk of 1/ 1,000

Note: Heavier weight grid cells denote data rather than test results
* Non-parametric one-tailed Mann-Whitney Wilcoxon test

Taken individually, the higher WTPs for children in Table 2 and the higher risk aversion for children in Table 3 both seem reasonable findings. However, the chaining approach combines these responses together in calculating VSLs and the child premium present in both value and risks seems to suggest the potential for double counting if chained together. We investigate the potential for such bias in the final study of this paper. However, for the moment we press on to the CV results from the present study.

(a) First response WTPs		Mean (£) *Median* (St. Dev.)		*Weak scope sensitivity test* p value (z statistic)*
		1/1000	*5/1000*	
Adult		441 *0* (1210)	448 *0* (1371)	0.267 (-0.62)
Child		1,047 *50* (3484)	4,187 *500* (18387)	<0.001 (-4.64)
*Adult vs child values**	p value (z statistic)	0.021 (-2.03)	<0.001 (-6.08)	

(b) Second response WTPs		Mean (£) *Median* (St. Dev.)		*Weak scope sensitivity test* p value (z statistic)*
		1/1000	*5/1000*	
Adult		458 *0* (1,728)	2941 *100* (18,451)	0.001 (-2.98)
Child		880 *25* (2,210)	2478 *100* (7,359)	0.063 (-1.53)
*Adult vs child values**	p value (z statistic)	0.003 (-2.78)	0.1014 (-1.27)	

(c) Comparing first & second response WTPs	*Are first and second responses different?* p value (z statistic)*	
	1/1000	*5/1000*
Adult	0.670 (0.44)	0.019 (-2.06)
Child	0.583 (0.21)	0.006 (-2.50)

Table 4: Summary statistics for WTP values from the CV responses (a) for first responses, (b) for second responses, and (c) comparing these

Note: Heavier weight grid cells denote data rather than test results

* Non-parametric one-tailed Mann-Whitney Wilcoxon test

Table 4 presents the results of the CV questions on WTP to avoid either a 1/1,000 or 5/1,000 risk of death to either the respondent or their child. To avoid the potential problems of anchoring, panel (a) of Table 4 only uses individuals' responses to the first of these four questions, the ordering of which were rotated across respondents as per Appendix 1. Results show clear differences in first response WTP for adult as opposed to child lives. Results from a weak scope sensitivity test (comparing WTP for a 1/1,000 risk to a 5/1,000 risk) show significant scope in child values but not in adult values which are dominated by zero WTP responses (medians for both risks being zero).

Panel (b) of Table 4 calculates the same values again but now using individuals' response to the second life they value. So, in an ordering which first asked a respondent to value a 1/1,000 risk reduction to their own life, panel (b) reports values based on that respondent's subsequent valuation of a risk reduction of 1/1,000 to their child's life (i.e. ordering Version 6 in Appendix 1). The potential for anchoring is obvious and reflected in findings. Now the adult valuations clearly pass a scope sensitivity because they have been upwardly anchored by their previous (relatively high) valuation of their child's life. Similarly valuations of child risk reductions are now anchored down by prior (relatively lower) values for adults. For example, considering the 5/1,000 risk reduction values for children, in panel (a) first responses provide a mean of £4,187 while in panel (b) anchoring on prior (low) adult) values gives a second response value of just £2,478 which is no longer significantly different from the adult value in that table. Consistency tests across first and second responses are reported in panel (c) which shows clear evidence of anchoring effects in CV values at the 5/1,000 risk level.

How do the chaining and CV methods compare? While almost none of the chaining respondents stated that they would not be willing to pay anything to avoid a given ill-health state[15], in the CV task the rate of zero WTP response ranged from 13% for the 5/1000 risk reduction for the child to 61% for the

[15] In the chaining exercise just over 1% of respondents gave a zero WTP to avoid a certain ill health state for themselves and no zero bids were recorded in respect of ill health states for children.

1/1000 reduction for the parent. High proportions of zero responses are a common and longstanding feature of the CV literature observed both in VSL studies (Krupnick et al., 2002)[16] and across a wide range of contexts and countries (see for example, Rowe and Chestnut, 1982; Desvousges et al., 1987; Bostedt and Boman, 1996; Jorgensen et al., 1999; Strazzera et al., 2003; Cho et al., 2008; Chen and Hua, 2015; Ferreira and Marques, 2015; Lee, 2015; Lee and Heo, 2016; Vossler and Holladay, 2018; and the recent review of the issue of zero responses in CV studies by Chen and Qi, 2018). Recalling that, within our field study, these are the same people who were happy to engage with the chaining exercise, the high rates of zero WTP and implied illogical zero VSL values, raise considerable concerns about the use of the CV method in this context, particularly where such problems are hidden by aggregate measures such as the mean.

Turning to consider those VSL estimates, following Krupnick et al. (2002)[17], for the CV data the VSL is calculated by taking the WTP for a particular change in the probability of death and dividing this by that probability change (Δp) as shown in Equation (1):

$$VSL = \frac{WTP_{\Delta p}}{\Delta p} \quad (1)$$

For the chaining method, and following the framework of Carthy et al. (1999) the stated WTP to avoid a particular ill-health state is coupled with the risk trade-off using Equation (2):

$$VSL = WTP \times \psi \quad (2)$$

where, to allow for potential framing effects, ψ is defined as $\psi_j = \frac{1-\delta}{\pi_j - \delta}$ where δ is the mortality risk associated with Treatment A (e.g. δ=1/1,000 in prior discussions), and π_j is the mortality risk level at which the respondent states she is indifferent between the Treatment A and Treatment B.

[16] Krupnick et al 2002 report a 36% rate of zero responses, a rate which lies in the middle of our observed range.

[17] A further approach is to use modelled rather than raw responses. However, this requires additional assumptions concerning the appropriate approach to modelling, assumptions which are somewhat contentious where the data is dominated by anomalous zero WTP responses, e.g. Krupnick et al. (2002) employ a spike model (Kriström, 1997). We prefer to avoid such assumptions and work with actual rather than modelled responses so as to adhere to the approach of Carthy et al. which is the main focus of our study.

		Mean (£) [Mean excluding zero] <5% Trimmed mean> *Median* *{Median excluding zero}* (St. Dev.)				Framing tests: Significance of anomalies	
Method		*CV*		*Chained*		*CV*[1]* *p value* *(z statistic)*	*Chained*[2]** *p value* *(z statistic)*
Risk level (CV) or Treatment (Chaining)		*1/1000*	*5/1000*	*X*	*Y*		
Adult		441,000 [1,143,333] <164,697> *0* *{500,000}* (1209864)	89,623 [199,484] <28,985> *0* *{30,000}* (274214)	11,377,436 [11,628,409] <1,289,649> *268,188* *{299879}* (87221816)	9,144,810 [9,211,561] <1,016,644> *138,119* *{151364}* (85107745)	0.424 (-0.19)	0.218 (-0.78)
Child		1,046,959 [1,936,875] <321,071> *50,000* *{500,000}* (3484135)	837,321 [965,048] <291,351> *100,000* *{200,000}* (3677320)	38,420,686 [38420686] <12,003,187> *1,125,857* *{1125857}* (139644139)	98,399,253 [98399253] <8,802,169> *2,002,004* *{2002004}* (853117844)	0.973 (1.93)	0.324 (-0.05)
*Adult vs child values**	*p value* *(z statistic)*	0.021 (-2.03)	<0.001 (-6.01)	<0.001 (-4.75)	<0.001 (-7.17)		

Table 5: VSL estimates using the chaining and CV approach

Note: 1. Significance of difference between CV VSL values based on WTP for 1/1,000 or 1/5,000 risk reductions

2. Significance of difference between Chained VSL values based on chaining from ill-health state X or Y

Heavier weight grid cells denote data rather than test results

* Non-parametric one-tailed Mann-Whitney Wilcoxon test; while standard theory expects no difference, the anomaly literature suggests that a VSL calculated from a 1/1000 risk may exceed that calculated from a 5/1000 risk.

** Non-parametric two-tailed Mann-Whitney Wilcoxon test; standard theory expects no difference and there is no indication of a directional effect from the anomaly literature.

The resulting VSLs from both the CV and chaining approaches are summarised in Table 5, which for the CV section uses only the first responses as previously described in panel (a) of Table 4 so as to mitigate against the anchoring which is a clear feature of this data. One result is very clear – given the means, medians, standard deviations and the fact responses are bound by zero, we observe highly positively skewed data.

Focussing upon the CV findings, here the skew is extreme with medians being zero for the adult values and relatively low for the child values. This results in mean values that are well below those given in the literature. In major part[18] this is likely to reflect the fact that we employ levels of risk similar to those observed in ordinary life (e.g. annual risks of car accidents, fatal cancers, etc. as per Viscusi, 1993) whereas the CV literature often uses much lower risk levels. The clear evidence of insensitivity to scope observed both in this paper, the wider literature and meta-analyses thereof (Lindhjem et al., 2011) strongly suggests that had we used small risk probabilities our CV derived VSL estimates would be significantly higher.

Untrimmed VSLs for the chaining method are particularly high for estimates of child values, suggesting that child premiums in both the valuation and MSG elements of the method may be causing a double counting bias; again we address this in our final study. Given that there are justifications for trimming extreme values (Chilton et al, 2015)[19], the chained trimmed mean values fall more in line with the extant literature although again the child VSL values remain somewhat but not implausibly high.

One clear message from Table 5 is that, within any method or starting point, VSL values for children are consistently higher than those for adults.

[18] See also our earlier footnote regarding our rejection of the Krupnick et al. decision to use modelled outputs from a spike model as opposed to raw WTP responses in calculating VSL. This is also likely to have avoided inflation of our CV VSL estimate.

[19] Chilton et al., note that a single high value would have resulted in "an untrimmed mean more than seven times higher than the median" (p. 297) in the Carthy et al., (1999) study had it not been removed from analysis.

Testing the internal consistency of the chaining method for estimating VSL: a second field study

Results from both our lab experiment and first field study point to significant issues for use of the CV method to estimate VSLs, most particularly in terms of anchoring and the insensitivity of WTP to changes in risk, especially where the absolute magnitude of the risks concerned are small, resulting to farming effects upon VSL estimates. In contrast, the robustness of the chaining method to such framing anomalies and very much lower rates of zero WTP responses (both problems being present in our CV results), gives some reason for cautious optimism regarding the usefulness of the approach. However, as outlined above, results from the first field study gave some cause for concern regarding the ability of the method to appropriately reflect any child premium expressed by parents. If parents use this premium to calculate both their WTP for (certainty) health improvements for their child, and also apply the same premium to their risk trade-offs on behalf of their child, then analysts' combination of these responses to generate VSL estimates may result in a double counting of this child premium.

The focus upon child versus parent VSL values allows us to propose a novel variant of the chaining consistency test devised by Carthy et al. The chaining approach to estimating VSL is to link together a single WTP valuation of avoiding a specified ill health state with a corresponding single risk trade-off question, linking that ill-health state to a risk of a fatal outcome. In our final field study we contrast such a 'single chain' approach with a 'double chain' variant of the chaining method. Here the risk trade-off is spilt into two parts; the first linking a minor (temporary) to a major (permanent) ill-health state, and the second linking that major ill-health state to mortality. If respondents are only incorporating their 'child premium' into the valuation element of the chaining process then a switch from the single to double chain variant should have no impact on resultant VSL. However, if that child premium is expressed at each stage of the chaining method then the double chain approach will

produce higher values than the single chain variant. Comparison of the single and double chain variants also allows us to test if the chaining approach is internally consistent more generally; adult VSLs should not vary significantly across these variants.

A further insight is provide by switching from the MSG to a more conventional Standard Gamble (SG) format (as discussed previously). Comparison of the single chain VSL obtained in this final study with that given in the previous study allows us to inspect the magnitude of any "certainty effect" within SG derived VSL values.

Questionnaire design

The questionnaire employed a customised Computer Aided Personal Interviewing (CAPI) program to visually communicate the risk probabilities[20] and easily randomise the order of treatments. Straightforward descriptions of the ill-health states were adopted using an approach similar to that of Baker et al. (2008) and yielding the following ill-health states (copies of the description cards seen by respondents are given in Appendix 2 while the questionnaire is available from the authors):

- *Temporary Illness Affecting Adult (Ta):* Severe stomach pains affecting the respondent with diarrhoea and vomiting for 2-3 days every 2 weeks for 12 months;
- *Temporary Illness Affecting Child (Tc):* Severe stomach pains affecting the respondents' child with diarrhoea and vomiting for 2-3 days every 2 weeks for 12 months;
- *Permanent Illness Affecting Adult (Pa):* Severe stomach pains affecting the respondent with diarrhoea and vomiting for 2-3 days every 2 weeks for the rest of life;
- *Permanent Illness Affecting Child (Pc):* Severe stomach pains affecting the respondents' child with diarrhoea and vomiting for 2-3 days every 2 weeks for the rest of life.

[20] The CAPI system conveyed risk probabilities both in terms of percentages and via a coloured grid similar to those used to convey risk in other stated preference studies (e.g. Zhang et al., 2013).

An initial, simple ranking exercise was used to raise respondents' understanding of these four ill-heath states[21]. All respondents were then asked to state their willingness-to-pay to avoid the certainty of each of the above illnesses (the first stage of the chaining procedure). WTP responses were elicited using a set of cards[22] (available from the authors) each detailing a separate payment amount which respondents sorted into categories ranging from 'definitely would pay' to 'definitely would not pay' with cards in intermediate (e.g. 'might pay') categories were then resorted until the maximum WTP was identified. The order in which cards were presented to respondents was randomised with the value on the first card being recorded to allow inspection of a potential starting point bias (Herriges and Shogren, 1996; Bateman et al., 2001).

Respondents were then presented with the risk trade-off questions variants using the SG format outlined previously. Respondents were asked to choose between the certainty of suffering one of the illness scenarios (*Ta*, *Tc*, *Pa*, *Pc* above) and an alternative risky treatment with some chance, 1-π_j, of complete recovery to full health and a π_j risk of a worse health outcome. This risky 'worse health state' was either: (1) the permanent condition (*Pa* or *Pc*) when considering the certainty of a temporary condition (*Ta* or *Tc*)[23] or (2) death when considering the certainty of a permanent condition (*Pa* or *Pc*). Starting from an initial level of π_j = 0.5 the CAPI varied this value according to the responses given until the respondent considered the risk of the worse health outcome as just equivalent to the certainty of the alternative health outcome. Respondents answered all permutations of these questions with the order of questions being randomised.

[21] Following best-practice guidance (e.g. Bateman et al., 2002; Champ et al., 2017) more than half of the sample completed a budget constraint task. Subsequent testing revealed that this had no significant impact upon subsequently stated WTP values.

[22] Cards ranged in value from £60 to £6,000,000 expressed as both lump-sum payments and as per month equivalents if costs were spread over ten years.

[23] The subject (adult or child) was kept constant at this stage. So if the adult was the subject of the permanent condition the adult would also be the subject of the temporary condition (and vice versa where the child was the subject.

Combining the WTP and SG responses allows us derive VSL estimates. The 'single chain' VSL is estimated using Equation 3:

$$\frac{WTP_p}{\pi_p} \tag{3}$$

While the 'double chain' approach estimates VSL using Equation 4:

$$\frac{WTP_t/\pi_t}{\pi_p} \tag{4}$$

The 'single chain' approach derives VSL by directly linking the willingness to pay to avoid the certainty of permanent illness (WTP_p) with that risk of death which the respondent feels is equivalent to the certainty of the permanent ill-health condition (π_p). The 'double chain' variant derives VSL more indirectly. First, we divide the willingness to pay to avoid the certainty of temporary illness (WTP_t) by that risk of permanent ill-health which the respondent feels is equivalent to the certainty of the temporary ill-health condition (π_t). This sum is then divided by that risk of death which the respondent feels is equivalent to the certainty of the permanent ill-health condition (π_p). Under standard theory, a respondents' WTP_P should be equal to WTP_t/π_t, hence the VSL values derived from the single and double chained variants should not differ significantly. Significant differences would suggest inconsistencies within the approach. Moreover, if any inconsistency is particularly apparent for child values then this would suggest that even the single chain VSL would not be robust to double counting.

The questionnaire concluded with a variety of socio-economic and follow-up questions, including the elicitation of respondents' household income[24].

Sampling procedure

Sampling was undertaking by a professional surveying company at a large number of locations across the UK to generate a representative sample of parents with children aged less than 18 years old. In total, 996 parents completed the survey. Tests confirm no significant difference in sub-samples across the various versions of the questionnaire.

Results

(a) Ranking and WTP to avoid illness

As per previous studies (Baker et al. 2008), the warm-up exercise showed that respondents generally rank impacts upon child health as more important that those affecting adults, with permanent impacts outranking temporary symptoms (details in Appendix 3).

Turning to consider respondents' WTP to avoid the four different ill-health states (*Pa*, *Pc*, *Ta*, *Tc*), very few respondents were observed to state a zero WTP value in the temporary illness valuations (1% for adult and 0.3% for child), and for permanent conditions all WTP values were strictly positive. These represent much lower proportions of zeroes than those typical of CV studies (see discussion of the first field study), a finding consistent with our first field study.

[24] Analysis of these variables showed that they did not materially affect the central results of this study and so they are excluded from further discussion.

Resultant WTP values are presented in Table 6 which is solely calculated from responses to the first ill-health state valued so as to avoid any possibility of ordering effects. Results conform to prior expectations with the values given to avoid a permanent ill-health condition always being significantly higher than those for avoiding a temporary condition and adult values being significantly lower than child values. As the temporary condition last for one year only, irrespective of the person affected, then results reflect a pure child premium. However, for the permanent condition this difference is exacerbated by the greater life expectancy of the child relative to the parent. This is reflected in the greater excess of mean WTP for children as opposed to adults in the permanent (as opposed to temporary) condition; relationships which appear to bolster the chained approach.

	Mean (£) *Median* (St. Dev.)		Are values to prevent permanent ill-health higher than those to prevent temporary ill-health?* p value (z statistic)
	Temporary ill-health	Permanent ill-health	
Adult	13,155 *3600* (46958)	27,766 *6000* (101906)	0.006 (2.50)
Child	18,354 *5999* (39289)	64,424 *9000* (427456)	0.001 (3.04)
Are child values higher than adult values?* p value (z statistic)	0.010 (2.32)	0.002 (2.83)	

Table 6: WTP to avoid the certainty of negative health outcomes
Note: Heavier weight grid cells denote data rather than test results
* Non-parametric one-tailed Mann-Whitney Wilcoxon test

The encouraging findings of Table 6 are tempered by those of Table 7 which report results from a regression analysis to test for starting point bias. This examines the effect on stated maximum WTP of the value displayed on the randomly selected first card shown to respondents. After controlling for the four ill-health states (*Pa*, *Pc*, *Ta*, *Tc*) we see a clear, positive and statistically significant relationship

between the amount shown on this first card and the final stated WTP. While such anchoring effects are common and long established in CV studies (Bateman et al., 1995; Green et al., 1998; Chien et al., 2005; Flachaire and Hollard, 2006) and indeed have been observed in incentivised, real payment experiments (Bateman et al., 2006), nevertheless the results of Table 7 suggest that the chaining method is not immune from such phenomena.

Predictor	Parameter (SE)	t value	p value
Intercept (*Pa*)	6.262 (0.209)	30.006	<0.001
Pc	0.369 (0.150)	2.470	0.011
Ta	-0.449 (0.175)	-2.564	0.014
Tc	-0.071 (0.175)	-0.404	0.686
Ln(starting bid)	0.301 (0.023)	13.137	<0.001

Table 7: results of a regression analysis testing for starting point bias
Notes: Dependent variable = natural logarithm of final stated (maximum) WTP
Adjusted R-squared (OLS estimator) = 0.175; N = 996.

(b) Standard Gamble results

Summary statistics of acceptable levels of risk (of the permanent condition when faced with a certainty of the temporary condition, and of death when faced with the certainty of the permanent condition) are reported in Table 8.

		Mean risk *Median* (St. Dev.)	
		Temporary	Permanent
Adult		0.212 *0.075* (0.266)	0.188 *0.065* (0.253)
Child		0.182 *0.075* (0.237)	0.132 *0.006* (0.229)
Adult vs child values*	p value (z statistic)	0.092 (-1.33)	0.003 (-2.79)

Table 8: Summary statistics for the risk values
Note: Heavier weight grid cells denote data rather than test results
* Non-parametric one-tailed Mann-Whitney Wilcoxon test

Table 8 reveals that parents are unwilling to allow their children to accept the same levels of risk that they would accept for themselves, with this difference in risk aversion being particularly significant for the more serious permanent ill-health state. Such results conform well to expectations and findings both in the health and other fields (Kahnemann and Tversky 1982; Jones-Lee, 1992; Gilovich and Medvec 1995; Connolly and Zeelenber 2002; Søgaard et al., 2012), however they suggest that respondents are applying a child premium in their risk responses, just as they did previously in the WTP questions. The implications for the chaining method of this double expression of a child premium are obvious and it is to these we now turn.

(c) VSLs for adult and child

As discussed earlier, the mean value for VSLs estimated through the chaining approach is highly susceptible to hyper-inflation by a few very extreme positive values. To combat this, one could

calculate a 'double-mean' ('double-median') value – using the mean (median) of the sample WTP and acceptable risk level values to arrive at the mean (median) VSL. Indeed, simply calculating this from the statistics in Tables 6 and 8 is very easy, and the results seem plausible if low in the case of adults[25]. However, this imposes a set of assumptions on societal preferences which are difficult to defend and yield values which are challenging to interpret. Instead, we calculate VSL at an individual level, using only the single or double chain first responses given by that individual (to minimise any ordering effect as respondents pass through the four, randomly ordered, ill-health states; *Pa*, *Pc*, *Ta*, *Tc*) and trimming the resultant data to remove the top and bottom 5%[26] of values to combat extremes. Table 9 reports the VSL findings which result from this procedure.

		Mean VSL (£million) *Median* *(St. Dev.)*		*Is the single chain value lower than the corresponding double value?** *p value* *(z statistic)*
		Single chain	Double chain	
Adult		1,743 0.22 (4,991)	332,747,600 3.69 (1,578,980,000)	<0.001 (-22.17)
Child		4,436 2.53 (11,069)	154,932,900 519.48 (5,958,649,000)	<0.001 (-23.24)
*Adult vs child values**	p value (z statistic)	<0.001 (-13.60)	<0.001 (-13.18)	

Table 9: Chained estimates of VSL for adults and children

Note: Calculated using Equations (3) and (4)
Heavier weight grid cells denote data rather than test results
* Non-parametric one-tailed Mann-Whitney Wilcoxon test

Ignoring the absolute values recorded in Table 9, at first glance these results appear promising. The VSL values for children are significantly larger than those for adults, conforming well to our

[25] Single chain VSL for the parent and child respectively; mean: £147,691 and £488,061; median: £92,308 and £1,500,000

[26] Chilton et al., (2015) defend the trimming of data. In the present case trimming reduces mean values by roughly one order of magnitude

expectations. However, comparing the mean single chain VSL with those estimated in the first field study, highlights that the "certainty effect" (induced by our switch from the MSG to SG format for the second field study) hyper-inflates estimated VSLs; respondents seem very unwilling to accept even small mortality risks when even a very adverse alternative ill-health state involves no risk of death. This inflation in resultant VSL is very substantially exacerbated when we move from the single to double chain format with the latter values being very significantly larger than the former. This exacerbation applies to both the child and adult values suggesting that not only does the method double count any child premium, it also double counts adults utility for their own health.

Concluding remarks

We present a set of lab and field exercises to examine the robustness of the CV and Chaining methods for estimating VSL values for both adults and children. Findings across these studies reveal a number of consistent results. The CV approach of asking respondents to value compound goods consisting of both risk levels and outcomes reveals a number of anomalies. Respondents struggle to comprehend risk levels giving inconsistent responses to the same probability levels expressed in different formats and over-valuing small as opposed to larger reductions in risk. The CV method also seems prone to rejection in the form of high rates of zero WTP bids for health risk reductions which cannot reasonably be described as having no value.

The chaining method therefore potentially offered an innovative response to the various problems exhibited by the CV approach, not least an unwillingness on the part of respondents to engage with such questions. The chaining method performs well in this respect with respondents appearing to understand and accept the constituent certainty valuation and risk trade-off elements of the method. However, the chaining approach appears just as vulnerable to starting point bias as does the CV

method. More uniquely the chaining approach seems vulnerable to an inflationary certainty effect when the risk trade-off is framed using conventional SG (as opposed to MSG) formats. Furthermore, and of most concern, when exposed to the test for internal consistency, the chaining approach clearly fails, double counting any premium to yield infeasibly high VSL values.

In short our study reveals substantial challenges to the application of both the CV and chaining methods. Given the vital importance of deriving robust VSL estimates for practical project appraisal and benefit cost analysis there is clearly considerable work to be done before these problems can be solved. However, both of our field studies using either method do reveal a very clear message to the policy question which prompted this investigation. We find strong evidence that parents place a higher VSL on their child that they do for their own lives.

Acknowledgements

Funding for this study was provided by the Valuation of Environment-Related Health Impacts: Accounting for Differences Across Age, Latency and Risk Categories with a Particular Focus on Children (VERHI Children) project, funded under CEC - Sixth Framework Programme grant FP6-2003-SSP-3. The funding body had no involvement in the study itself, or in the decision to submit the article. We are grateful to steering group representatives from both the US EPA and the EEA (including Richard Carson, Maureen Cropper, James Hammitt and Alan Krupnick) whose advice we gratefully acknowledge. We also owe a great debt to Graham Loomes for his input throughout this research and to Andreas Isoni for discussion of findings. All opinions expressed in this paper and errors therein are the responsibility of the authors.

References

Agee, M., and T. Crocker. (1996). Parental Altruism and Child Lead Exposure: Inferences from the Demand for Chelation Therapy, *The Journal of Human Resources*, 31(3): 677-691.

Alberini, A. (2005). What Is a Life Worth? Robustness of VSL Values from Contingent Valuation Surveys, *Risk Analysis*, 25(4): 783-800.

Alberini, A. and Ščasný, M. (2011). Context and the VSL: Evidence from a Stated Preference Study in Italy and the Czech Republic, *Environmental and Resource Economics*, 49(4): 511–538.

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

Alolayan, M. A., Evans, J. S., and Hammitt, J. K. (2017). Valuing mortality risk in Kuwait: Stated-preference with a new consistency test. *Environmental and Resource Economics*, *66*(4), 629-646.

Baker, R., Bateman, I., Donaldson, C., Jones-Lee, M., Lancsar E., Loomes, G., Mason H., Odejar M., Prades J-L. P., Robinson A., Ryan, M., Shackley P., Smith, R., Sugden, R. and Wildman J. (2008). Weighting and valuing quality adjusted life years: preliminary results from the Social Value of a QALY Project. Report for the Methodology Programme

Baker, R. and Robinson, A. (2004). Responses to Standard Gambles: Are Preferences 'Well Constructed'? *Health Economics,* 13(1): 37-48.

Bateman, I.J. Cole, M. Cooper, P. Georgiou, S. Hadley, D. and Poe, G.L. (2004) On visible choice sets and scope sensitivity, *Journal of Environmental Economics and Management*, 47, 71-93

Bateman, I.J., Carson, R.T., Day, B., Hanemann, W.M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Özdemiroğlu, E., Pearce, D.W., Sugden, R. and Swanson, J. (2002) *Economic Valuation with Stated Preference Techniques: A Manual*, Edward Elgar Publishing, Cheltenham.

Bateman, I.J., Langford, I.H., Jones, A.P. and Kerr, G.N. (2001) Bound and path effects in multiple-bound dichotomous choice contingent valuation, *Resource and Energy Economics*, 23(3): 191-213.

Bateman, I.J., Langford, I.H., Turner, R.K., Willis, K.G. and Garrod, G.D. (1995) Elicitation and truncation effects in contingent valuation studies, *Ecological Economics*, 12(2):161-179. DOI: 10.1016/0921-8009(94)00044-V

Bateman I J, Munro A, Rhodes B, Starmer C and Sugden R. (2006) Anchoring and yea-saying with private goods: an experiment, in List, J.A. (Ed.) *Using Experimental Methods in Environmental and Resource Economics*, Edward Elgar, Cheltenham, pp. 1-19.

Beattie, J., Covey, J., Dolan, P., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., Robinson, A., and Spencer, A. (1998) On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 1—Caveat Investigator, *Journal of Risk and Uncertainty*, 17, 5–25.

Bellavance, F., Dionne, G., & Lebeau, M. (2009). The value of a statistical life: A meta-analysis with a mixed effects regression model. *Journal of Health Economics, 28*(2), 444–464.

Blomquist G.C., Dickie M., O'Conor R.M., (2011). Willingness to pay for improving fatality risks and asthma symptoms: Values for children and adults of all ages. *Resource and Energy Economics*, Resource and Energy Economics, Volume 33, Issue 2, May 2011, Pages 410-425.

Bostedt, G., and M. Boman (1996). Non-response in Contingent Valuation-Reducing Uncertainty in Value Inference, *Environmental and Resource Economics,* 8(1): 119-124.

Braathena, N-A., Lindhjem H., and Navrud S., (2008) Valuing Lives Saved From Environmental, Transport and Health Policies: A Meta-Analysis of Stated Preference Studies. Working paper, OECD Headquarters, Paris.

Brown, T. T. (2015). The Subjective Well-Being Method of Valuation: An Application to General Health Status. *Health services research*, *50*(6), 1996-2018.

Carson, R.T. (2011) *Contingent Valuation: A Comprehensive Bibliography and History*, Edward Elgar, Cheltenham and New York.

Carson, R.T. and Groves, T. (2007) Incentive and informational properties of preference questions, *Environmental and Resource Economics*, 37(1): 181-120, doi.org/10.1007/s10640-007-9124-5

Carthy, T., S M Chilton, J Covey, L Hopkins, M W Jones-Lee, N Pidgeon and A Spencer, (1999). The Contingent Valuation of Safety and the Safety of Contingent Valuation, Part 2: The CV/SG 'Chained' Approach, *Journal of Risk and Uncertainty*, 17: 187-213.

Champ, P. A., & Bishop, R. C. (2001). Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias. *Environmental and Resource Economics*, *19*(4), 383-402.

Champ, P.A., Boyle, Kevin and Brown, Tom C. (eds.) (2017) *A Primer on Non-market Valuation, 2nd Edition*, The Economics of Non-Market Goods and Services: Volume 15, Springer, Dordrecht.

Chen, B. and Qi, X. (2018) Protest response and contingent valuation of an urban forest park in Fuzhou City, China, *Urban Forestry & Urban Greening*, 29: 68-76, doi.org/10.1016/j.ufug.2017.11.005.

Chen, W.Y. and Hua, J. (2015) Citizens' distrust of government and their protest responses in a contingent valuation study of urban heritage trees in Guangzhou, China, *Journal of Environmental Management*, 155: 40-48, doi.org/10.1016/j.jenvman.2015.03.002.

Chien, Y., Hauang, C.J. and Shaw, D. (2005) A general model of starting point bias in double bounded dichotomous contingent valuation survey, *Journal of Environmental Economics and Management*, 50: 362–377.

Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N.F., Robinson, A. and Spencer, A. (2004) Valuing the 'value' of life: a question of constructed preference? in Todd, Z., Nerlich B., McKeown S. and Clarke D. (Eds.) *Mixing Methods in Psychology*, Routledge, London.

Chilton, S., Covey, J., Jones-Lee, M.W., Loomes, G., Pidgeon, N. and Spencer, A. (2015) Response to 'Testing the validity of the "value of a prevented fatality" (VPF) used to assess UK safety measures', *Process Safety and Environmental Protection*, 93: 293-298, doi.org/10.1016/j.psep.2014.11.002.

Cho, S., Yen, S., Bowker, J., & Newman, D. (2008). Modeling Willingness to Pay for Land Conservation Easements: Treatment of Zero and Protest Bids and Application and Policy Implications. *Journal of Agricultural and Applied Economics*, 40(1), 267-285. doi:10.1017/S1074070800028108

Connolly, T. and Zeelenber, M. (2002) Regret in Decision Making. *Current Directions in Psychological Science*, 11(6): 212 -216.

Cropper, C., Hammitt, J.K. and Robinson, L.A. (2011) Valuing Mortality Risk Reductions: Progress and Challenges, *RFF DP 11-10*, Resources For the Future, Washington, D.C.

Desaigues, B., Ami, D., Bartczak, A., Braun-Kohlová, M., Chilton, S., Czajkowski, M., Farreras, V., Hunt, A., Hutchison, M., Jeanrenaud, C., Kaderjak, P., Máca, V., Markiewicz, O., Markowska, A., Metcalf, H., Navrud, S., Nielsen, J.S., Ortiz, R., Pellegrini, S., Rabl, A., Riera, R., Scasny, M., Stoeckel, M-E., Szántó, R. and Urban, J. (2011) Economic valuation of air pollution mortality: A 9-country contingent valuation survey of value of a life year (VOLY) (2011) *Ecological Indicators*, 11(3): 902-910, doi.org/10.1016/j.ecolind.2010.12.006.

Desvousges, W.H., Johnson, F.R., Dunford, R.W., Hudson, S.P., Wilson, K.N., Boyle, K.J. (1992). Measuring Nonuse Damages Using Contingent Valuation: An Experimental Evaluation of Accuracy. Research Triangle Institute Monograph 92-1. Research Triangle Park, NC.

Desvousges, W.H., Smith, V.K. and Fisher, A. (1987) Option price estimates for water quality improvements: A contingent valuation study for the monongahela river, *Journal of Environmental Economics and Management*, 14(3): 248-267, doi.org/10.1016/0095-0696(87)90019-2.

DfT (Department for Transport) (2016) *Transport Analysis Guidance data book*, available at www.gov.uk/government/publications/webtag-tag-data-book-july-2016.

Dickie, M., and Gerking, S. (2003). Valuation of Environmental Risks to Children's Health. *Working paper*, Department of Economics University of Central Florida, Orlando, FL 32826.

Dolan, P., Metcalfe, R., Munro, V., and Christensen, M.C. (2008). Valuing lives and life years: Anomalies, implications, and an alternative. *Health Economics*, Policy and Law, 3(3), 277-300. doi:10.1017/S1744133108004507

Donaldson, C., Baker, R., Mason, H., Jones-Lee, M., Lancsar, E., Wildman, J., Bateman, I.J., Loomes, G., Robinson, A., Sugden, R., Pinto Prades, J-L., Ryan, M., Shackley, P. and Smith, R. (2011) The social value

of a QALY: raising the bar or barring the raise?, *BMC Health Services Research*, 11:8, http://www.biomedcentral.com/1472-6963/11/8

Dubourg, W.R., Jones-Lee, M.W. and Loomes, G. (1997) Imprecise Preferences and Survey Design in Contingent Valuation, *Economica*, 64, 681–702.

European Union, (2001). *Recommended Interim Values for the Value of Preventing a Fatality in DG Environment Cost Benefit Analysis*, available from http://ec.europa.eu/environment/enveco/others/recommended_interim_values.pdf.

Ferreira, S. and Marques, R.C. (2015) Contingent valuation method applied to waste management, *Resources, Conservation and Recycling*, 99: 111-117, doi.org/10.1016/j.resconrec.2015.02.013.

Flachaire, E. and Hollard, G. (2006) Controlling starting point bias in double-bounded contingent valuation survey, *Land Economics*, 82, 103–111.

Gilovich, T., & Medvec, V. H. (1995). The experience of regret: What, when, and why. *Psychological Review*, 102(2), 379.

Green, D., Jacowitz, K.E., Kahneman, D. and McFadden, D. (1998) Referendum contingent valuation, anchoring, and willingness to pay for public goods, *Resource and Energy Economics*, 20(2): 85-116, doi.org/10.1016/S0928-7655(97)00031-6.

Guerriero, C., Cairns, J., Bianchi, F., & Cori, L. (2018). Are children rational decision makers when they are asked to value their own health? A contingent valuation study conducted with children and their parents. *Health economics*, *27*(2), e55-e68.

Gyrd-Hansen, D., Kjær, T., & Nielsen, J. S. (2012). Scope insensitivity in contingent valuation studies of health care services: should we ask twice? *Health economics*, *21*(2), 101-112.

Gyrd-Hansen, D., Jensen, M. L., & Kjaer, T. (2014). Framing the willingness-to-pay question: Impact on response patterns and mean willingness to pay. *Health economics*, *23*(5), 550-563.

Hammar, H. and Johansson-Stenman, O. (2004) The value of risk-free cigarettes – do smokers underestimate the risk? *Health Economics*, 13: 59-71. doi:10.1002/hec.794

Hammitt, J.K. and J.D. Graham (1999) Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability, *Journal of Risk and Uncertainty*, 8; 33-62.

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

Hammitt, J.K., Zhou, Y. (2006). The Economic Value of Air-Pollution-Related Health Risks in China: A Contingent Valuation Study, *Environmental and Resource Economics,* 33(3): 399-423.

Hara Associates Inc. (2000). *Benefit/cost analysis of proposed tobacco products information regulations*. Prepared for Health Canada and Consulting and Audit Canada. Ottawa, Ontario, June 5, 2000.

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

Herriges, J.A. and Shogren, J.F. (1996) Starting Point Bias in Dichotomous Choice Valuation with Follow-Up Questioning, *Journal of Environmental Economics and Management*, 30(1): 112-131, https://doi.org/10.1006/jeem.1996.0008.

Hunt, A., Ortiz R. A. (2006). Review of Revealed Preference Studies. Paper presented at the VERHI-Children Advisory Group Meeting, OECD, Paris, November 2006.

Jacowitz, K.E. and Kahneman, D. (1995) Measures of Anchoring in Estimation Tasks, *Personality and Social Psychology Bulletin*, 21(11): 1161 – 1166, doi.org/10.1177/01461672952111004

Jenkins, R., Owens, N., and Wiggins, L. (2001). Valuing Reduced Risks to Children: The Case of Bicycle Safety Helmets, *Contemporary Economic Policy,* 19(4): 397-408.

Jones-Lee, M.W. (1989) *The Economics of Safety and Physical Risk*, Basil Blackwell, Oxford.

Jones-Lee, M.W., (1992) Paternalistic Altruism and the Value of Statistical Life, *Economic Journal*, 102(410), 80-90.

Jones-Lee, M.W., and Loomes, G., (2004) Eliciting Measures of Value for Health and Safety. 9th Australian Workshop on Safety-Related Programmable Systems (SCS'04), Brisbane. *Conferences in Research and Practice in Information Technology*, Vol. 47.

Jones-Lee, M.W., Loomes, G., and Philips, P. (1995) Valuing the Prevention of Nonfatal Road Injuries: Contingent Valuation vs Standard Gambles, *Oxford Economic Papers*, 47:676-695.

Jorgensen, B.S., Syme, G.J., Bishop, B.J. and Nancarrow, B.E. (1999) Protest Responses in Contingent Valuation, *Environmental and Resource Economics*, 14(1): 131–150, https://doi.org/10.1023/A:1008372522243

Kahneman, D., and Tversky, A. (1982) The psychology of preferences. *Scientific American*, 246: 160-173.

Kriström, B. (1997) Spike Models in Contingent Valuation, *American Journal of Agricultural Economics*, 79(3): 1013–1023, https://doi.org/10.2307/1244440

Krupnick, A., Alberini, A., Cropper, M., Simon, N., O'Brien, B., Goeree, R. and Heintzelman, M. (2002). Age, Health, and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents, *Journal of Risk and Uncertainty*, (24): 161-184.

Lindhjem, H., Navrud, S., Braathen, N.A. and Biausque, V. (2011) Valuing mortality risk reductions from environmental, transport, and health policies: a global meta-analysis of stated preference studies, *Risk Analysis*, 31(9):1381-407, doi: 10.1111/j.1539-6924.2011.01694.x.

Lee, J-S. (2015) Measuring the benefits of the Intangible Cultural Heritage Hall in Jeonju Korea: Results of a Contingent Valuation Survey, *Journal of Cultural Heritage*, 16(2): 236-238, doi.org/10.1016/j.culher.2014.05.001.

Lee, C-Y and Heo, H. (2016) Estimating willingness to pay for renewable energy in South Korea using the contingent valuation method, *Energy Policy*, 94: 150-156, doi.org/10.1016/j.enpol.2016.03.051.

McDaid, D., Sassi, F., & Mekrur, S. (2015). Supporting effective and efficient policies: the role of economic analysis. In Promoting Health, Preventing Disease. Open University Press.

McFadden, D., & Train, K. (Eds.). (2017). *Contingent Valuation of Environmental Goods: A Comprehensive Critique*. Edward Elgar Publishing.

Morgan, M.G. (2017) *Theory and Practice in Policy Analysis: Including Applications in Science and Technology*, Cambridge University Press, Cambridge.

Mount, T., W. Schulze, W. Weng, N. Zhang, and L. Chestnut. (2003) The Effects of Age and Family Status on the Value of Statistical Life: Evidence from the Automobile Market and a National Survey of Automobile Use. Presented at US EPA's workshop "Valuing Environmental Health Risks to Children", October 2003, Washington DC.

Nastis, S., and T. Crocker (2003). Pregnant Mother's Valuation of Own and of Child Health. Presented at the USEPA Workshop on Valuing Environmental Health Risks to Children, Washington, DC, October 20-21, 2003.

Nielsen, J. S., Gyrd-Hansen, D., & Kjær, T. (2012). Valuation of morbidity and mortality risk reductions. Does context matter? *Accident Analysis & Prevention*, *48*, 246-253.

O'Brien, J. (2013) The Age-Adjusted Value of a Statistical Life: Evidence from Vehicle Choice, Georgetown University, available at https://pdfs.semanticscholar.org

Pinto-Prades, J-L., Martinez-Perez, J-E. and Abellán-Perpiñán, J-M. (2006) The influence of the ratio bias phenomenon on the elicitation of health states utilities, *Judgment and Decision Making*, 1(2): 118–133, http://journal.sjdm.org/jdm06126.pdf

Peters, E., N. Dieckmann, D. Västfjäll and C. Mertz (2005a). When Five out of Four People have Trouble with Fractions and Other Numbers: Numeracy and Mood in Decisions. University of Oregon.

Peters, E., P. Slovic and J. Hibbard (2005b). Bringing Meaning to Numbers: Evaluability and Affect in Choice. University of Oregon, Eugene, Oregon.

Prades, P. Jones-Lee, M.W., Loomes G. and Brey, R., (2009). Trying to estimate a monetary value for the QALY. Working paper, University Pablo de Olavide.

Rakotonarivo, O.S., Schaafsma, M. and Hockley, N. (2016) A systematic review of the reliability and validity of discrete choice experiments in valuing non-market environmental goods, *Journal of Environmental Management*, 183(1): 98-109, doi:10.1016/j.jenvman.2016.08.032.

Rice, T. (2015). Reflecting on 'Valuing lives and life years: anomalies, implications, and an alternative', *Health Economics, Policy and Law*, *10*(4), 405.

Robinson, A., Gyrd-Hansen, D., Bacon, P., Baker, R., Pennington, M., Donaldson, C., & Team, E. (2013). Estimating a WTP-based value of a QALY: the 'chained' approach. *Social Science & Medicine*, *92*, 92-104.

Roldós, M. I., Corso, P., & Ingels, J. (2017). How much are Ecuadorians Willing to Pay to Reduce Maternal Mortality? Results from a Pilot Study on Contingent Valuation. *International journal of MCH and AIDS*, *6*(1), 1.

Rowe, R.D. and Chestnut, L.G. (1982) *The value of visibility: Economic theory and applications for air pollution control*, Abt Books, Cambridge, Mass.

Scotton, C. R. (2013) New risk rates, inter-industry differentials and the magnitude of VSL estimates. *Journal of Benefit-Cost Analysis*, *4*(1), 39-80.

Shepard, D.S. and Zeckhauser, R.J. (1984) Survival versus consumption, *Management Science*, 30(4): 423–439.

Shogren, J.F. (1990) The impact of self-protection and self-insurance on individual response to risk, *Journal of Risk and Uncertainty*, 3(2): 191–204, doi.org/10.1007/BF00056372

Slovic, P. (1995). The Construction of Preference, *American Psychologist*, 50: 364-371.

Smith, V. K. and W. H. Desvousges (1987). An Empirical Analysis of the Economic Value of Risk Changes, *Journal of Political Economy,* 95: 89-114.

Søgaard, R., Lindholt, J., & Gyrd-Hansen, D. (2012). Insensitivity to Scope in Contingent Valuation Studies. *Applied health economics and health policy*, *10*(6), 397-405.

Strazzera, E., Scarpa, R., Calia, P., Garrod, G.D. and Willis, K.G. (2003) Modelling zero values and protest responses in contingent valuation surveys, *Applied Economics*, 35:2, 133-138, DOI: 10.1080/0003684022000015900

Thomas, P.J. and Vaughan, G.L. (2015) Testing the validity of the "value of a prevented fatality" (VPF) used to assess UK safety measures, *Process Safety and Environmental Protection*, 94: 239-261, doi.org/10.1016/j.psep.2014.07.001.

Tversky, A. and D. Kahneman, (1973). Availability: A Heuristic for Judging Frequency and Probability, *Cognitive Psychology*, 5: 207-232.

Tversky, A. and D. Kahneman (1974). Judgement under Uncertainty: Heuristics and Biases, *Science*, 185: 1124-1131.

Viscusi, W.K. (1993) The Value of Risks to Life and Health, *Journal of Economic Literature*, 31(4): 1912-1946, www.jstor.org/stable/2728331

Viscusi, W. K. (2009) The devaluation of life, *Regulation & Governance*, *3*(2), 103-127.

Vossler, C.A. and Holladay, J.S. (2018) Alternative value elicitation formats in contingent valuation: Mechanism design and convergent validity, *Journal of Public Economics*, 165: 133-145, doi.org/10.1016/j.jpubeco.2018.07.004.

Zhang, J., Adamowicz, W., Dupont, D.P. and Krupnick, A. (2013) Assessing the extent of altruism in the valuation of community drinking water quality improvements, *Water Resources Research*, 49, 6286–6297, doi: 10.1002/wrcr.20447.

Appendix 1: The eight treatments of the first field survey

Version 1	Version 2	Version 3	Version 4
Chaining Method	**Chaining Method**	**Chaining Method**	**Chaining Method**
1. WTP to prevent ill-health X to parent	1. WTP ill-health Y to parent	1. WTP ill-health X to child	1. WTP ill-health Y to child
2. MSG linking X to death for parent	2. MSG Y for parent	2. MSG X for child	2. MSG Y for child
3. WTP to prevent ill-health Y to child	3. WTP ill-health X to child	3. WTP ill-health Y to parent	3. WTP ill-health X to parent
4. MSG linking Y to death for child	4. MSG X for child	4. MSG gamble Y for parent	4. MSG X for parent
Direct WTP 10 year risk	**Direct WTP 10 year risk**	**Direct WTP 10 year risk**	**Direct WTP 10 year risk**
5. WTP to reduce mortality risk by 5/1000 for parent	5. WTP 5/1000 parent	5. WTP 5/1000 child	5. WTP 5/1000 child
6. WTP to reduce mortality risk by 1/1000 for parent	6. WTP 1/1000 parent	6. WTP 1/1000 child	6. WTP 1/1000 child
7. WTP to reduce mortality risk by 5/1000 for child	7. WTP 5/1000 child	7. WTP 5/1000 parent	7. WTP 5/1000 parent
8. WTP to reduce mortality risk by 1/1000 for child	8. WTP 1/1000 child	8. WTP 1/1000 parent	8. WTP 1/1000 parent
Version 5	**Version 6**	**Version 7**	**Version 8**
Chaining Method	**Chaining Method**	**Chaining Method**	**Chaining Method**
1. WTP ill-health X to parent	1. WTP ill-health Y to parent	1. WTP ill-health X to child	1. WTP ill-health Y to child
2. MSG X for parent	2. MSG Y for parent	2. MSG X for child	2. MSG Y for child
3. WTP ill-health Y to child	3. WTP ill-health X to child	3. WTP ill-health Y to parent	3. WTP ill-health X to parent
4. MSG Y for child	4. MSG X for child	4. MSG gamble Y for parent	4. MSG X for parent
Direct WTP 10 year risk	**Direct WTP 10 year risk**	**Direct WTP 10 year risk**	**Direct WTP 10 year risk**
5. WTP 1/1000 parent	5. WTP 1/1000 parent	5. WTP 1/1000 child	5. WTP 1/1000 child
6. WTP 5/1000 parent	6. WTP 5/1000 parent	6. WTP 5/1000 child	6. WTP 5/1000 child
7. WTP 1/1000 child	7. WTP 1/1000 child	7. WTP 1/1000 parent	7. WTP 1/1000 parent
8. WTP 5/1000 child	8. WTP 5/1000 child	8. WTP 5/1000 parent	8. WTP 5/1000 parent

Appendix 2: Illness card descriptors handed to respondents

Ta

WHO IS AFFECTED	YOU
SYPTOMS	SEVERE STOMACH PAINS, DIARRHOEA AND VOMITING FOR 2-3 DAYS EVERY 2 WEEKS
LENGTH OF ILLNESS	12 MONTHS

Tc

WHO IS AFFECTED	YOUR CHILD
SYPTOMS	SEVERE STOMACH PAINS, DIARRHOEA AND VOMITING FOR 2-3 DAYS EVERY 2 WEEKS
LENGTH OF ILLNESS	12 MONTHS

Pa

WHO IS AFFECTED	YOU
SYPTOMS	SEVERE STOMACH PAINS, DIARRHOEA AND VOMITING FOR 2-3 DAYS EVERY 2 WEEKS
LENGTH OF ILLNESS	THE REST OF YOUR LIFE

Pc

WHO IS AFFECTED	YOUR CHILD
SYPTOMS	SEVERE STOMACH PAINS, DIARRHOEA AND VOMITING FOR 2-3 DAYS EVERY 2 WEEKS
LENGTH OF ILLNESS	THE REST OF YOUR CHILDS LIFE

Appendix 3: The ranking of illness scenarios in the second field study

	Perceived Severity			
Symptom	I (highest)	II	III	IV (lowest)
$Permanent_{child}$	57%	31%	10%	2%
$Permanent_{adult}$	38%	42%	16%	4%
$Temporary_{child}$	2%	18%	47%	33%
$Temporary_{adult}$	2%	10%	27%	61%