Published as: Andersson, H., Hole, A.R., and Svensson, M., "Valuation of small and multiple health risks: A critical analysis of SP data applied to food and water safety", Journal of Environmental Economics and Management, vol. 75, January 2016, pp. 41–53.

Valuation of small and multiple health risks: A critical analysis of SP data applied to food and water safety

[Running title: Valuation of small and multiple health risks]

Henrik Andersson*, Toulouse School of Economics (LERNA, UT1C, CNRS), France

Arne Risa Hole, University of Sheffield, United Kingdom

Mikael Svensson, Karlstad University, and Health Metrics, The Sahlgrenska Academy, University of Gothenburg, Sweden

Abstract: This study elicits individual risk preferences in the context of an infectious disease using choice experiments. A main objective is to examine scope sensitivity using a novel approach. Our results suggest that the value of a mortality risk reduction (VSL) is highly sensitive to the survey design. Our result cast doubt on the standard scope sensitivity tests in choice experiments, but also on the validity and reliability of VSL estimates based on stated-preference studies in general. This is important due to the large empirical literature on non-market evaluation and the elicited values' central role in policy making.

Keywords: Choice experiments; Morbidity risk; Mortality risk; Scope sensitivity; Time preferences; Willingness to pay

JEL-Codes: D61; H41; I18 ; Q51

1. Introduction

The use of benefit-cost analysis (BCA) to guide resource allocation in environmental, health and risk policies that affect mortality and morbidity requires a common metric for benefits and costs. Monetary values act as this common metric and today there is broad consensus that the willingness to pay (WTP) approach to evaluate health risk reductions, which was established in the 1960s and early 1970s (Dreze 1962, Jones-Lee 1974, Mishan 1971, Schelling 1968), is the appropriate approach to evaluate small changes in health risks. Since the early theoretical contributions a vast amount of empirical work has been conducted, evaluating a wide range of risks (Viscusi and Aldy 2003, Lindhjelm et al. 2011). Whereas there is consensus about the WTP approach there has been controversy regarding the empirical elicitation of individual WTP.

Since no easily available market prices exist for health risk reductions, researchers instead rely on either revealed- (RP) or stated-preference (SP) methods. RP methods use individuals' actual decisions in markets that are related to the good of interest, whereas SP methods elicit individuals' preferences in constructed and hypothetical markets using surveys. The empirical literature on monetizing health risks has mostly dealt with the WTP to reduce mortality risks, referred to as the value of a statistical life (VSL). The literature has been dominated by the hedonic pricing technique (Rosen 1974), an RP method which uses market data, for instance in the areas of workplace and traffic safety (Viscusi and Aldy 2003, Andersson and Treich 2011), and the SP method referred to as the contingent valuation method (CVM) (Lindhjelm et al. 2011). In recent years discrete choice experiments (DCE) have become a popular SP approach to evaluate health risks in health, transportation, and environmental economics (Hensher et al. 2009, Cameron, DeShazo, and Stiffler 2010, Alberini et al. 2007, Adamowicz et al. 2011). Carson and Louviere (2011) defined DCE as an SP method where respondents are asked to make a discrete choice between two or more options in a choice set, and where the different options are varied based on an experimental design. The results from the DCE studies on health risk evaluation are in line with results from both RP and CVM studies; individuals have a positive WTP to reduce their risk exposure, WTP varies between contexts, and the population means of mortality and morbidity risks are similar to values from the other evaluation techniques. Reviews and meta-analyses of the WTP literature on mortality health risk evaluation have shown that most VSL estimates fall within the range US$ 1 to 10 million (Dekker et al. 2011, Lindhjelm et al. 2011, Viscusi and Aldy 2003).[1]

Among economists not much controversy has surrounded the empirical application of the RP methods; since actual decisions are used individuals have incentives to be well informed and to make decisions that are in their interest. More controversy has surrounded the validity of SP approaches (CVM and DCE), generally based on the fact that decisions are hypothetical (see e.g. Hausman 2012). Despite the criticism there has been a large increase in the use of SP studies over the past few decades (Carson and Hanemann 2005). The advantage of the SP approach is that it offers flexibility in creating specific markets of interests and allows the analysts to control the decision alternatives. Advocates of SP studies have argued that much of the early negative findings were a result of bad survey design and that more recent surveys, where the methodology and knowledge among analysts have improved, provide results with a much stronger validity (Hammitt and Haninger 2010, Carson 2012).

---

[1] Ranges and reference years for the price levels varies between studies. However, the US$ 1 to 10 million range is in line with the narrower range of the estimates using workplace safety data reported in Viscusi and Aldy (2003). A topic not addressed thoroughly in these cited reviews is whether SP studies systematically produce higher or lower estimates than RP studies. Reviews of the literature have come to different conclusions, though, with, e.g. de Blaeij et al. (2003) finding evidence that SP studies produce higher estimates than RP studies and Kochi et al. (2006) the opposite.

To evaluate the validity of the WTP estimates from SP studies analysts rely to a large extent on scope sensitivity tests (Arrow et al. 1993, Desvousges, Mathews, and Train 2012), i.e. that WTP increases with the scope of the good being valued. Researchers can examine scope sensitivity using a within-sample ("internal") and/or between-sample ("external"/"split-sample") test. In a within-sample test each respondent answers a number of valuation questions with varying scope of the good, whereas in a between-sample test different sub-samples value different scopes of the good. The within-sample test is far less demanding since by asking repeated questions (provide repeated choice sets) respondents are given a clue to show "internal consistency" (Charness, Gneezy, and Kuhn 2012, Fischhoff and Frederick 1998). In the literature on WTP for health risk reductions, a distinction is also made between "weak" and "strong" sensitivity to scope (Corso, Hammitt, and Graham 2001). Weak scope sensitivity is fulfilled if WTP increases with the size of the risk reduction, while strong scope sensitivity refers to the situation where WTP increases near-proportionally to the magnitude of the risk reduction (Hammitt 2000). A lack of near-proportionality will result in a large variation in the estimates of the marginal WTP for the risk reduction, which will question the validity of the results (Hammitt 2000).The general finding in the literature is that WTP is sensitive to the size of the good but not in line with what theory predicts (Hammitt and Graham 1999, Desvousges, Mathews, and Train 2012, Lindhjelm et al. 2011), i.e. "weak" scope is usually satisfied, whereas "strong" scope is not.

The CVM has been extensively evaluated using both within- and between-sample tests, whereas DCEs almost exclusively have been evaluated using within-sample tests. By construct, each DCE provides a within-sample scope sensitivity test considering that the scope of the good is varied across choice-sets, and significant scope is thus often described to be present if the coefficient estimate of an attribute (e.g. the size of the mortality risk reduction) show the expected sign and is statistically significant. Some authors have argued that the between-sample test does not apply to DCEs (Adamowicz et al. 2011). In a meta-analysis of VSL estimates Lindhjelm et al. (2011) found that 291 of 318 estimates passed a within-sample scope test, whereas only 85 of 199 CVM estimates passed a between-sample scope test. The authors did not examine whether the choice of CVM or DCE affected the sensitivity to scope, but only CVM estimates were evaluated using the more demanding between-sample test. In the wider literature, Desvousges, Mathews, and Train (2012) conducted a survey of CV studies on natural resource damage assessment and found that 40 of 109 studies passed all scope tests, the rest showed mixed results or failed to show sensitivity to scope. Studies evaluating relative sensitivity to scope in CVM vs. DCE have also used between-sample (CVM) to within-sample (DCE) comparisons (see e.g. Foster and Mourato 2003). Already the NOAA panel (Arrow et al. 1993) was explicit about the fact that between-sample scope sensitivity tests are necessary to demonstrate valid estimates from an SP study. The lack of evaluations of DCE studies in terms of between-sample scope sensitivity, or something of the sort, is therefore a fundamental missing piece in the literature.[2]

Thus, in this study we extend the literature on DCEs and scope sensitivity tests by performing both the standard within-sample validity test, as well as a novel between-sample scope sensitivity test. The novelty of our approach to examining this question is that we for one subsample run a state-of-the-art design where we use the actual policy-relevant mortality risk-reduction levels and then for another sample use levels that are significantly higher, but still reasonable from the respondents' perspective. This makes it possible to test for scope sensitivity both within each sub-sample (whether respondents within a sub-sample prefer policies with larger risk reductions) and between sub-samples (whether respondents in a sub-sample with larger risk reductions report a higher WTP).

We address our aim using an application of health risks related to an infectious disease caused by the bacteria campylobacter, i.e. campylobacteriosis. Humans are mainly infected by campylobacter through contaminated food or water (Taylor et al. 2012) and we will therefore elicit preferences in a market setting where individuals can reduce their risk by consuming safer food or water. The data are modelled using both standard and latent-class logit models with non-linear utility functions to allow for the estimation of discount rates. The latent class model allows for preference heterogeneity, which we find is substantial, and makes it possible to investigate differences among groups of respondents in terms of their sensitivity to scope.

The paper is structured as follows. In section 2 we describe our data collection and show some descriptive statistics. Section 3 shows our econometric approach outlining the conditional logit and latent

---

[2] It should be mentioned that there is a literature which investigates price vector effects in DCEs (e.g. Slothuus Skjoldborg and Gyrd-Hansen 2003, Hanley, Adamowicz, and Wright 2005, Carlsson and Martinsson 2008) . Using between-sample designs these studies find that changing the range and levels of the price attribute can have an impact on estimates of marginal WTP. These studies do not investigate sensitivity to scope in the context of risk reductions, however.

class models, whereas results are shown in section 4. Section 5 concludes the paper with a discussion of the results and their place in the literature.

2. The Survey and Data Collection

In the DCE respondents were asked to choose between different public policies that were described to reduce campylobacter-related mortality and morbidity risks. One concern about SP studies in general has been that their hypothetical nature may be perceived to be inconsequential by the respondents (Carson and Groves 2007). Respondents were told when invited to take part in the survey the information that they will provide is important for how society allocates resources, and we assume that respondents treated the survey as consequential. The policies differed across choice sets with respect to the size of mortality and morbidity risk reductions, the source of the disease being targeted (food- or water-borne), when the policy would start to have an effect, and the monetary cost of the policy. To address the issue of within- and between-sample scope sensitivity we created a split-sample design with two sub-samples that were identical in all aspects with the exception of the size of the mortality risk reduction. We refer to the two sub-samples as sub-sample A and sub-sample B, where sub-sample A was exposed to the vector of smaller risk-reductions.

2.1 Survey Structure

Following an introductory welcome note to respondents, the survey consisted of four sections. The first section contained questions on respondents' risk perception and attitudes towards food and water safety, personal experience of food poisoning as well as a set of questions regarding respondents' risk behavior (e.g. their use of risk-reducing measures in the home environment). Section two described the illness of campylobacteriosis to the respondents. The annual incidence was described to be 63 000 in Sweden, which corresponds to a risk of 7 in 1,000 (AgriFood 2012). It was further described that campylobacteriosis can be categorized as mild, moderate or severe with accompanying symptoms described. In section two the respondents were also asked to state their health status using a visual analog scale. Section three contained the DCE where respondents were asked to choose between policies (or the status quo alternative) that differed with respect to the levels of the respective attributes. Following the DCE part, the fourth section included questions on socio-economics and demographics. After the fourth section, respondents could choose to finalize their participation in the survey, but they were also asked if they would consider answering a number of debriefing questions.

In order to design the survey in a comprehensible and clear way we initially tested the survey in small focus groups. Following this, we performed two pilot studies on-line with 100 and 50 respondents. The feedback from the two pilot studies induced some minor textual changes to the description of the risk scenario and some modifications of attribute levels.

2.2 Attributes and Levels

The choice experiment was designed with 5 attributes with a varying number of attribute levels: source of disease (2 levels), mortality risk reduction (3 levels), morbidity risk reduction (3 levels), delay (4 levels), and cost (3 levels). Table 1 below shows the attributes and their levels.

[Table 1 about here]

The levels of each attribute were determined based on relevance to the research question, discussions and feedback from a medical expert in the field of infectious diseases, as well as feedback from the focus groups. The first attribute listed in Table 1 is the source of disease, i.e. food- or water-born campylobacteriosis. This attribute will be irrelevant to respondents if they only care about the size of the risk reduction. But if respondents perceive there to be different averting behavior possibilities, or consider the controllability or dreadfulness etc. of the risks to differ, the source of disease can affect risk preferences (Slovic 2000, Shogren and Crocker 1999). Here we hypothesize that perceived controllability of the risk is lower for water-borne campylobacter and that this may positively affect the valuation of water-borne risk reducing policies.

To address our main research question we take a novel approach and design two alternative scenarios varying the mortality risk vector. In the first scenario (sub-sample A) the mortality risk reductions varied between 1, 2 and 4 fewer deaths per year. These mortality risk reductions are reasonable and policy-relevant given the current number of deaths due to campylobacteriosis in Sweden, which was reported to the respondents of sub-sample A as less than five cases among the 63 000 people becoming sick every year. In the second scenario (sub-sample B) the mortality risk-reductions were multiplied by a factor of 100 to be in line with road-fatality risk. In sub-sample B it was also explained that 63,000 people get sick every year, but the number of deaths due to campylobacteriosis each year was not specifically mentioned. It was only stated that in rare events the illness can lead to death.[3] The motivation for using risk reductions in line with road-fatality risk is that there is a large body of empirical evidence, not only internationally (Andersson and Treich 2011) but also based on Swedish data (Hultkrantz and Svensson 2012) based on these risk levels. Hence, this provides an opportunity to compare our estimates to previously reported values in the literature. This meant, however, that the objective baseline mortality-risk could not be communicated to the respondents in sub-sample B since it is smaller than the risk reductions used in the choice sets.

By running the analysis on sub-sample A and B, respectively, we can examine within-sample scope sensitivity by examining the size and significance of the coefficient estimate of the mortality risk reduction attribute; i.e. the standard scope sensitivity test in the DCE literature. By comparing estimates between sub-sample A and B, we can also, conduct a test analogous to a between-sample scope sensitivity test. Not only does this design permit us to test for within- and between-scope sensitivity, but it also makes it possible to relate our estimates based on sub-sample B with the VSL literature in previous Swedish and other international studies (Hultkrantz and Svensson 2012, Lindhjelm et al. 2011).

The levels for the morbidity risk reductions were chosen as to represent sizeable effects and to be in balance with mortality risk reductions, i.e. neither of the attributes would obviously dominate the other. The levels were discussed in focus groups and established in the pilot surveys (they were initially slightly lower).

To help respondents understand the small changes in risk we took a number of different actions. Due to the very small mortality risk-levels we use frequencies rather than probabilities in the risk description in order to lessen the cognitive burden for respondents and also making it possible for respondents to relate the risk numbers to the total population at risk, i.e. the Swedish population (Kalman and Royston 1997). As

---

[3] In sub-sample A the information respondents received after being informed about the risk of illness was "In addition, in rare events the illness can lead to death (less than 5 cases in total per year in Sweden).". Since this baseline risk is smaller than the risk reductions used in sub-sample B, and since providing another baseline level would have meant stating an incorrect baseline risk, the information about the number of cases was left out in sub-sample B. Based on debriefing questions we have no indications that respondents in sub-sample B considered the attribute levels as less reasonable or relevant compared to respondents in sub-sample A.

discussed by Slovic, Monahan, and MacGregor (2000), presenting risk in frequencies rather than probabilities may make it easier for respondents to show scope sensitivity. Moreover, for the morbidity risk we used a verbal probability analog by describing that the total number of illnesses correspond to 700 cases in city of 100 000 habitants (an average sized city in Sweden). Due to the very low baseline mortality-risk in sub-sample A, and that no explicit baseline mortality-risk was provided in sub-sample B, together with the mixed evidence found in the literature on visual aids' ability to solve the scope sensitivity issue (e.g. Corso, Hammitt, and Graham 2001, Haninger and Hammitt 2011, Goldberg and Roosen 2007), we decided to not include any visual aid. Instead, in addition to the measures mentioned above, after the first choice set the respondents received feedback on their choice in terms of what it implied in changes in costs, prevented deaths, prevented illnesses, etc. The respondent was then asked if he/she was satisfied with his/her choice and wanted to proceed to the next choice set, or to change his/her choice in the current choice set. This procedure was provided to enhance the understanding of the risk scenario and choice alternatives for the respondents, especially considering that interactive feedback on choices/decisions generally has been shown to be influential for learning and understanding (Hattie and Timperley 2007).

The attribute delay, reflecting when the beneficial effect of the policy would start to have an effect varied between 0, 2, 5 and 10 years. The cost of the project would be immediate for the respondent, i.e. the delay only concerns when the benefits will have effect. The levels for the cost attribute were determined partly to cover reasonable ranges for respondents' budget set, but also to allow for a large range of possible estimates of VSL as well as for the value of a statistical illness (VSI) (Lindhjelm et al. 2011), and finally adjusted based on the results from the pilot studies.

On the basis of all possible combinations in the full factorial design, 64 choice sets with two alternatives were constructed using a D-optimal design algorithm (Carlsson and Martinsson 2003). The 64 choice sets were blocked into eight versions, and each respondent was faced with eight choice sets. The order of the choice sets was randomized across respondents.

2.3 The Choice Sets

Before the respondents were faced with the choice sets a general description of the policy scenario was stated as (freely translated from Swedish):

"[A]ssume that a government authority is considering two different policies that can reduce the occurrence of campylobacter; a stricter food control or improved water sanitation. We are interested in your valuation of these policies and will now ask you to answer 8 different questions. Apart from the fact that the policies differ with respect to the focus on food or water-spread campylobacter, the policies also differ regarding: the number of fewer deaths, the number of fewer illnesses, when the policy starts to have a beneficial effect and the cost of the policy. Once the policies become effective the risk reduction will last for 5 years, after that the risk of campylobacteriosis returns to its original level. Regardless of when risk reduction becomes effective the payment for the policy starts today and continues for 5 years."[4]

To increase realism the policies were described to have a 5-year duration. However, the risk and cost attributes were presented as annual changes to make them directly comparable and easier to interpret and understand by the respondents. Figure 1 shows an example of a choice set that was faced by sub-sample A, where respondents prior to answering the first choice set were reminded of the baseline risk. As shown, the respondents were asked to choose between two different policies (Policy A or Policy B) or choosing the status quo alternative, i.e. preferring to have neither of the policies implemented.

(Today 5 people die and 63 000 people get sick every year due to Campylobacterios. We now ask you to state if you prefer a certain policy (or not) to reduce these risks for a given cost. What do you prefer?)

The text in the parenthesis at the top of the figure was only presented to the respondents in the first choice set. The only difference in sub-sample B was that the number of deaths per year was not stated and risk reduction levels were multiplied by 100, everything else was identical to choice-sets in sub-sample A.

[Figure 1 about here]

---

[4] In order to make sure that elicited preferences reflect the health and cost domains as stated in the choice experiment it was further explained to respondents that the social insurance system would compensate potential income losses and health care costs. We also included a "cheap talk script" in order to mitigate some of the potential hypothetical bias that may arise in SP studies. The payment mechanism was generic for all policies in order to not have a confounding effect of different payment mechanisms on the interpretation of our results.

As reported above, after the respondent's first choice he/she was provided feedback on the computer screen on the meaning of his/her choice and after this respondents had the possibility to change their initial choice. We found that 16.8 percent of the respondents changed their initial choice. In the following 7 choice sets respondents were not given the possibility to change their decisions and they could not click ahead before responding to the current choice set (i.e. reading ahead was not possible).

2.4 Data

The data collection took place during the spring of 2012 and was conducted on-line using a web-panel of respondents (conducted by the company Scandinfo). Respondents were recruited to the web-panel by phone (there was no "self-recruitment" to the panel) at random among internet-enabled individuals in Sweden aged 18 and over. This does not necessarily mean that it constitutes a random sample of all Swedish citizens, but considering that Sweden has among the highest Internet penetration rates in the world (ITU 2012) it is a region where it may be made a strong case for using a web-based study. In total 1 250 respondents were included, where 1 000 respondents were randomly selected into sub-sample A and 250 respondents were randomly selected into sub-sample B.

Table 2 shows descriptive statistics for sex, age, university education, employment and income for sub-sample A and B together with a comparison to national population statistics for Sweden (SCB 2011, 2010). There are no statistically significant differences between sub-sample A and B for any of the background variables in the data. In comparison with national statistics our sample corresponds well or quite well with regards to sex, age, employment and income. It corresponds less well with the share of individuals with a university education (3 years or more); with 32-34 percent of our sample having a university education compared to 19 percent in the Swedish population (in the age range 18+).

[Table 2 about here]

Table 3 shows descriptive statistics of the sample regarding risk experience, risk perception, and subjective knowledge of food and water poisoning. Once again we see that there are no statistically significant differences across sub-sample A and B. Eight and 12 percent of the sub-samples report to have been food poisoned during the last year, whereas (in both sub-samples) eight percent report to have been food poisoned due to campylobacter (ever, i.e. not only during the last year). On average, the respondents in both sub-samples perceive the risk of being food poisoned (during a year) to be larger compared both to the incidence of food poisoning reported among our respondents and to objective national statistics. Whereas the objective annual risk of food poisoning is in the order of 10 per 100, the average perceptions among the respondents are 16.73 to 17.27 per 100. This however is the arithmetic mean. The geometric mean, which is common to use when analyzing risk perception since it reduces the effect from outliers (Viscusi and Hakes 2004, Andersson 2011, Andersson and Lundborg 2007), is 10.40 and 9.88 for sub-sample A and B, respectively, and not statistically significantly different from the objective risk. Also regarding the perceptions of the individual risk of being food poisoned due to campylobacter the arithmetic means suggest that respondents perceive their risk to be above objective average risks; 16.35 to 25.65 per 1 000 compared to objective risks of 7 per 1 000, but again the geometric means suggest the opposite, 3.85 and 4.03. Finally in Table 3 we report data on the respondents' self-assessed health using a visual analog scale ranging from 0 to 100 (with 100 representing "perfect health"), with mean responses at 80.09 and 81.94 (levels in line with previous Swedish findings (Brooks et al. 1991, Koltowska-Häggström et al. 2007, Andersson et al. 2013, Andersson and Lundborg 2007)).

[Table 3 about here]

3. Empirical model

3.1 Baseline model

As described in the previous section the individuals who participated in the experiment were asked to choose their preferred option out of a total of *J*=3 alternatives (two hypothetical scenarios and the status-quo) in *T*=8 choice sets. In our baseline specification the utility that respondent *n* derives from choosing alternative *j* in choice set *t* is given by

$$U_{njt} = sq + \beta_1 die_{njt} \exp(-\delta delay_{njt}) + \beta_2 sick_{njt} \exp(-\delta delay_{njt}) + \beta_3 water_{njt} + \beta_4 cost_{njt} + \varepsilon_{njt} \qquad (1)$$

where $\beta_1,...,\beta_4$ are coefficients to be estimated, sq is an alternative-specific constant for the status quo alternative and ε_{njt} is a random error term which is assumed to be IID type I extreme value. Assuming constant exponential discounting $die_{njt}\exp(-\delta delay_{njt})$ and $sick_{njt}\exp(-\delta delay_{njt})$ represent the discounted reductions in the risk of death and illness, respectively, where δ is the discount rate for mortality and morbidity, respectively (Alberini and Ščasný 2011). The remaining attributes in the utility function are described in Table 1.

The increase in cost necessary to keep the utility of an individual unchanged following the introduction of a policy which lowers the probability of dying without delay is given by

$$-\frac{\partial U_{njt} / \partial die_{njt}}{\partial U_{njt} / \partial cost_{njt}} = -\frac{\beta_1}{\beta_4} \tag{2}$$

This is a measure of the VSL since it can be interpreted as the WTP for a reduction in risk equivalent to saving one life. By replacing the variable *die* with *sick*, we get the VSI, which can be interpreted as the WTP for a reduction in risk equivalent to preventing one case of campylobacteriosis.

3.2 Latent class model

The baseline specification assumes that the respondents have identical preferences for the attributes of the policies, which is unlikely to be the case in reality. We explore this by estimating latent-class models, in which the utility function is given by

$$U_{njt} = sq_c + \beta_{c1} die_{njt}\exp(-\delta_c delay_{njt}) + \beta_{c2} sick_{njt}\exp(-\delta_c delay_{njt}) + \beta_{c3} water_{njt} + \beta_{c4} cost_{njt} + \varepsilon_{njt} \tag{3}$$

The subscript *c*, where $c = 1,...,C$, indicates the class membership of the individual respondent. The latent class model extends the standard logit model by allowing the preferences of respondents in different classes to vary, while maintaining the assumption of preference homogeneity within classes. A further advantage of the latent class model is that it takes the panel structure of the data into account.

Conditional on membership in class *c* the probability that respondent *n* chooses alternative *j* in choice set *t* is

$$L_{njt|c} = \frac{\exp(V_{njt|c})}{\sum_{j=1}^{J}\exp(V_{njt|c})} \tag{4}$$

where $V_{njt|c}$ is the deterministic (non-random) part of the utility function (Train 2009). Following Hensher and Greene (2003) we specify the probability that respondent *n* belongs to class *c* as

$$H_{nc} = \frac{\exp(\gamma_c' Z_n)}{\sum_{c=1}^{C} \exp(\gamma_c' Z_n)} \tag{5}$$

where Z_n is a vector of characteristics relating to individual *n* and γ_C is normalised to zero for identification purposes. In the application we set Z_n=1, which implies that the class membership probabilities are constant across respondents.

Combining equations 4 and 5 the unconditional probability of respondent *n*'s sequence of choices is given by

$$P_n = \sum_{c=1}^{C} H_{nc} \prod_{t=1}^{T} \prod_{j=1}^{J} (L_{njt|c})^{y_{njt}} \tag{6}$$

where y_{njt} is 1 if respondent *n* chose alternative *j* in choice set *t* and 0 otherwise. In the baseline case where there is only one class this model reduces to the standard conditional logit model. The parameters in the model are estimated by maximizing the log-likelihood function

$$LL = \sum_{n=1}^{N} \ln P_n \tag{7}$$

4. Results

4.1 Baseline model results

Table 4 presents the result of the baseline model estimated on sub-sample A and B. It can be seen that respondents prefer policies with lower costs and which lead to greater reductions in the probability of death and illness. Hence, in both sub-sample A and B the typical (weak) within-sample scope sensitivity test is satisfied. The insignificant status-quo constant suggests that the average respondent does not have a preference for or against the status-quo alternative holding the other attributes constant.[5] There are no qualitative differences between the sub-samples in terms of the sign and significance of the coefficients.

[Table 4 about here]

As explained in section 4 the coefficients in the utility function can be used to derive estimates of VSL and this is where the difference between the two samples becomes apparent, which can also be seen in Table 4. According to the model estimated on sub-sample A the VSL is SEK 4 732 million (95% CI: 3954–5509) (USD 710 million), while according to the model estimated on sub-sample B the VSL is SEK 70 million (95% CI 45-95) (USD 11 million). Hence, with a 100 times smaller risk reduction in sub-sample A the VSL is about 68 times larger, which implies that the sensitivity to scope is limited. Using the complete combinatorial approach described in Poe, Giraud, and Loomis (2005) we can reject the null hypothesis that the sub-sample A VSL is 100 times larger at the 5% level, but not at the 1% level. This suggests that there is, at best, very weak evidence of sensitivity to scope.[6]

In contrast, the VSI estimates are identical in the two sub-samples at SEK 0.49 million (approx. USD 0.07 million). Hence, when changing the mortality risk reduction between the two-sub samples, we get large effects on estimated VSL whereas we get no effect on the VSI where there is no change in the risk reduction. Figure 2 plots the coefficient ratios[7] in the model estimated on subsample B against the corresponding coefficient ratios in the model estimated on subsample A. It can be seen that with the exception of mortality risk the coefficient ratios in the two models are similar.

[Figure 2 about here]

Regarding other results we find that respondents have a preference for policies that are water rather than food-based, and that come into effect sooner rather than later. The estimated discount rates are very

---

[5] This result holds whether we use dummy or effects coding (Bech and Gyrd-Hansen 2005). We have used dummy coding in the reported models.

[6] We carried out a further test of weak sensitivity to scope by examining if a higher proportion of respondents choose the alternative with the greatest mortality reduction in sub-sample B. While we do find that the average proportion is higher in sub-sample B (49.8% vs. 44.4%), the difference is relatively small. Moreover, we find that in 16 (27%) of the 59 choice sets where the mortality risk attribute differs the proportion is actually higher in sub-sample A. This supports our conclusion that the evidence for scope sensitivity is, at best, very weak.

[7] The coefficients have been divided by the negative of the cost coefficient to eliminate differences in scale across the models. The *sick* coefficient has been multiplied by 10,000 to have a comparable magnitude to the other coefficients.

similar in the two subsamples, and in the order of 9-10%. Early studies to empirically estimate discount rates related to health varied substantially (Frederick, Loewenstein, and O'Donoghue 2002). However, many recent studies have found discount rates in line with our estimates, i.e. in the range 7-14% (Alberini et al. 2007, Viscusi, Huber, and Bell 2008, Rheinberger 2011, Meyer 2013).[8]

4.2 Latent class model results

Table 5 presents the result of latent class models with 2 classes estimated on the sub-sample A and B.[9] The latent class results suggest that there are two groups of respondents with markedly different preferences for the attributes in the experiment. In both the sub-sample A and B models there is a majority class of respondents who have a relatively high sensitivity to risk reductions, as reflected in high estimates of VSL and VSI. These respondents also have a negative and significant status quo constant, suggesting that they prefer to introduce a policy rather than maintaining the status quo, all else equal. Conversely there is a class of respondents who have a low sensitivity to risk reductions and a positive and significant status quo constant, which can be taken as evidence that they prefer to keep the status quo unless the benefits of the proposed policy are large.[10]

Importantly, although there are large differences between the estimated VSL and VSI of the two classes of respondents in each sub-sample, the estimated VSL in the “low-VSL” class in sub-sample A of SEK 1 186 million (95% CI: 590–1782) (USD 178 million) is an order of magnitude greater than the estimated VSL in the “high-VSL” class in sub-sample B of SEK 89 million (95% CI: 65-113) (USD 13 million). This suggests that preference heterogeneity cannot explain the large difference in average VSL observed across the sub-samples, which strengthens our conclusion that the difference is due to insensitivity to scope.

In terms of discount rates there are no big differences between the classes in sub-sample A, with both groups of respondents having an estimated discount rate of about 12%. The estimated discount rate

---

[8] These estimates are higher than discount rates applied by public authorities in health and safety decision making which typically lies in the region 1.5-3.5% (Robinson and Hammitt 2011, NICE 2011, Quinet, Baumstark, and al. 2013, ASEK 2014).

[9] The models were estimated using the *optimize* Stata/Mata function and code written by the authors. Several different sets of starting values were used to minimise the chance of the algorithm getting trapped in a local maximum. It was also attempted to estimate models with more than two classes, but these models did not converge.

[10] In sub-sample B the estimated VSL and VSI for this latter group of respondents are both insignificantly different from zero, suggesting that these respondents prefer to keep the status quo regardless of any reduction in risk.

in the “high VSL” class in sub-sample B is also about 12%, while the corresponding estimate for the “low-VSL” class is about 19%. It should be noted that this latter estimate is imprecise; we cannot reject the null hypothesis that the estimated discount rate in the “low-VSL” class is equal to 0.

The estimates of VSL and VSI averaged over classes using the class membership probabilities as weights are similar to the baseline estimates. The average VSL estimate in the model estimated on sub-sample A is SEK 4 636 million (95% CI: 4053–5218) (USD 695 million) compared to SEK 69 million (95% CI: 50–87) (USD 10 million) in subsample B. As in the baseline model the VSI estimates are virtually identical in the two sub-samples at SEK 0.56-0.57 million (approx. USD 0.08 million).

[Table 5 about here]

[Table 6 about here]

5. Discussion

This study employed a discrete choice experiment to elicit preferences for food and water safety related to campylobacteriosis. The major objective was to analyze within- and between-sample scope sensitivity of the respondents’ WTP for risk reductions. To answer the research question we constructed a DCE study with attributes including different levels of both mortality and morbidity risk reductions, where respondents were randomized to one of two sub-samples that differed (by a factor of 100) in the attribute levels of the mortality risk-reduction.

The results from our baseline model within the two sub-samples are in line with expectations; respondents prefer the benefits of the program sooner than later, programs that reduce both the mortality and morbidity risk, and less costly programs. Moreover, our results suggest that respondents prefer water-compared with food-safety programs (everything else equal), which is in line with the hypothesis that water risk is less controllable than food risk, and hence, WTP is higher for the former. When extending our analysis with latent class models we find clear evidence of preference heterogeneity. In particular, we find a majority group of respondents who have a relatively high sensitivity to risk reductions, as well as a minority group who have a much lower sensitivity to risk reductions. There is evidence of within-sample scope-sensitivity (i.e. significant risk reduction coefficients with the expected sign) in both models, and the results thus pass the typical within-sample scope test as implemented in most DCE studies. When we compare the results between samples, however, we find clear evidence of insensitivity to scope. In our baseline logit model the estimated VSL was SEK 4 732 million (95% CI: 3954–5509) (USD 710 million) and SEK 70 million (95% CI 45-95) (USD 11 million) in sub-samples A and B, whereas the results in the latent-class model (averaged across classes) were SEK 4 636 million (95% CI: 4053–5218) (USD 695 million) and SEK 69 million (95% CI: 50–87) (USD 10 million) in sub-sample A and B. Hence, with a 100 times smaller risk reduction in sub-sample A the VSL is about 67 times larger.

In a recent meta-analysis containing 850 estimates VSL was shown to vary between USD 4 450 and USD 197 million with a weighted mean VSL at USD 7.4 million (Lindhjelm et al. 2011). While our results from sub-sample A are outside of this range our sub-sample B results fall very well within the range of previous published estimates, and are of a similar magnitude to the reported weighted mean in the meta-analysis. Moreover, the estimates are in line with previous published estimates of VSL related to road safety in Sweden that in a recent review were shown to vary between USD 0.7 and 8.3 million with a mean and median equal to USD 2.9 and 2 million (Hultkrantz and Svensson 2012). This finding is of interest since the mortality risk level in sub-sample B was based on the risk levels for road-mortality risk in Sweden.

The between-sample analysis suggests extremely limited scope sensitivity, which questions the validity of our estimates as reflecting respondents' "true preferences". Our estimates of VSI are, however, robust between our sub-samples which we expected since the risk reductions did not change between the sub-samples. Note, though, that this is not evidence that our VSI is a valid estimate of respondents' preferences. If the morbidity risk reductions also had been altered between sub-samples, we may have experienced the same scope insensitivity as for the mortality risk. Moreover, and more importantly from a methodological perspective, the between-sample comparison highlights that even if a study finds within-sample scope sensitivity, which is the typical scope test in DCE studies, this does not necessarily suggests that the estimated WTP is a valid measure of individual preferences. As discussed by e.g. Goldberg and Rosen (2007) the systematic and repeated questions respondents answer in the DCE approach may stimulate a desire of respondents to be "internally consistent", i.e. respondents anchor their decisions on early choices and in subsequent choice sets to a larger degree state to prefer policies with larger risk reductions (and lower prices). This “coherent arbitrariness” creates a pattern in the data that will lead to a rejection of weak scope insensitivity within samples but not necessarily across samples using different scopes of the risk reduction (Ariely, Loewenstein, and Prelec 2003, Charness, Gneezy, and Kuhn 2012, Fischhoff and Frederick 1998), precisely what we find in our study.

Our results add to the broad and extensive literature on the validity of SP studies in general, where much focus has been placed on the issue of scope (in)sensitivity (Kahneman, Ritov, and Schkade 1999, Kahneman and Knetsch 1992, Hausman 2012, Carson 2012). Already in the blue ribbon panel convened by NOAA it was stated that scope insensitivity constitutes “perhaps the most important internal argument against the reliability of the CV approach” (Arrow et al. 1993, p.4607). Some authors have argued that scope insensitivity is avoidable in well conducted SP studies and has highlighted that insensitivity to scope has been rejected in many studies (Carson 1997, 2012) and further that it is something also observed in individuals’ behavior in some real market transactions (Randall and Hoehn 1996). Others have argued that scope insensitivity is likely to prevail in SP studies irrespective of survey design quality, due to concerns such as answers to a large extent reflecting “moral satisfaction” or expression of attitudes rather than economic WTP for the good/program or attribute (Kahneman and Knetsch 1992, Kahneman, Ritov, and Schkade 1999). If preferences are constructed in the survey setting, rather than pre-existing and well-defined, as suggested by e.g. Payne et al. (2000), it is unsurprising that analysts find within-sample, but not between-sample, scope sensitivity.

Irrespective of the strength of the different arguments, when it comes to the application of valuing mortality risk reductions, lack of near-proportional scope sensitivity (which is almost never found) undermines the results and implies very large variances in actual estimates of VSL. And even proponents of SP methods (see e.g. Carson 2012) highlight that one area of application that seems to be particularly prone to scope insensitivity is valuing changes in small probabilities. Findings from the psychological literature also highlight that individuals’ tend to not use numeric information when evaluating alternatives with small absolute differences and when the domain is unfamiliar, i.e. as is the case for most people when evaluating mortality risks (Peters et al. 2009).

To conclude, by carrying out a between-sample test for sensitivity to scope we have shown that the within-sample tests commonly used in DCEs are not sufficient to test the validity of respondents' WTP. In line with evidence from the CVM literature we find that the more demanding between-sample test suggests the existence of, at best, weak sensitivity to scope. Moreover, our findings indicate that obtaining estimates in line with other studies, which is often taken as evidence that the estimates are reliable, can be a result of analysts using similar methodologies, based on the same or similar risk scenario, and conducted in the same geographical area. Thus, standard reliability tests, which examine whether the estimates are in line with other findings in the literature, may be misleading. The finding that DCEs are susceptible to issues of insensitivity to scope comparable to those documented in the CVM literature raises the question of which of the two SP methods is better suited to elicit individual preferences for small risk changes. Our study does not address this question, but it is an important area for future research given the relevance of non-market evaluation methods for resource allocation.

References

Adamowicz, Wiktor, Diane Dupont, A Krupnick, and Jing Zhang. 2011. "Valuation of cancer and microbial disease risk reductions in municipal drinking water: An analysis of risk context using multiple valuation methods." *Journal of Environmental Economics and Management* 61:213-226.

AgriFood. 2012. "Policy Brief 2010:2." *AgriFood: economics centre, Lund, Sweden.*

Alberini, A, Stefania Tonin, Margherita Turvani, and Aline Chiabai. 2007. "Paying for permanence: Public preferences for contaminated site cleanup." *Journal of Risk and Uncertainty* 34 (2):155-178.

Alberini, Anna, and Milan Ščasný. 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.

Andersson, Henrik. 2011. "Perception of Own Death Risk: An Assessment of Road-Traffic Mortality Risk." *Risk Analysis* 31 (7):1069-1082.

Andersson, Henrik, James K. Hammitt, Gunnar Lindberg, and Kristian Sundström. 2013. "Willingness to Pay and Sensitivity to Time Framing: A Theoretical Analysis and an Application on Car Safety." *Environmental and Resource Economics* 56 (3):437-456.

Andersson, Henrik, and Petter Lundborg. 2007. "Perception of own death risk: An analysis of road-traffic and overall mortality risks." *Journal of Risk and Uncertainty* 34 (1):67-84.

Andersson, Henrik, and Nicolas Treich. 2011. "The Value of a Statistical Life." In *The Handbook in Transport Economics*, edited by A. de Palma, R. Lindsey, E. Quinet and R. Vickerman, 396-424. Cheltenham, UK: Edward Elgar Publishing Ltd.

Ariely, Dan, George Loewenstein, and Drazen Prelec. 2003. "Coherent Arbitrariness: Stable Demand Curves without Stable Preferences." *Quarterly Journal of Economics* 118 (1):73-105.

Arrow, Kenneth J., Robert Solow, Paul Portney, E. Leamer, R. Radner, and H. Schuman. 1993. "Report of the NOAA Panel on Contingent Valuation." *58 Federal Register 46*.

ASEK. 2014. Samhällsekonomiska principer och kalkylvärden för transportsektorn: ASEK 5.1 - Kapitel 22 English summary of ASEK Guidelines. Borlänge, Sweden: Trafikverket.

Bech, Mickael, and Dorte Gyrd-Hansen. 2005. "Effects coding in discrete choice experiments." *Health Economics* 14:1079-1083.

Brooks, Richard G., Stefan Jendteg, Björn Lindgren, Ulf Persson, and Stefan Björk. 1991. "EuroQol: Health-Realted Quality of Life Measurement. Results of the Swedish Questionnaire Exercise." *Health Policy* 18:37-48.

Cameron, Trudy Ann, J.R. DeShazo, and Peter Stiffler. 2010. "Demand for health risk reductions: A cross-national comparison between the U.S. and Canada." *Journal of Risk and Uncertainty* 41 (3):245-273.

Carlsson, F., and P. Martinsson. 2008. "How much is too much? An investigation of the effect of the number of choice sets, context dependence and the choice of bid vectors in choice experiments." *Environmental and Resource Economics* 40 (2):165-176. doi: 10.1007/s10640-007-9146-z.

Carlsson, Fredrik, and Peter Martinsson. 2003. "Design techniques for stated preference methods in health economics." *Health Economics* 12:281-294.

Carson, Richard T. 1997. "Contingent Valuation and Tests of Insensitivity to Scope." In *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*, edited by R.J. Kopp, W. Pommerhene and N. Schwartz, 127-163. Boston: Kluwer.

Carson, Richard T. 2012. "Contingent Valuation: A Practical Alternative when Prices Aren't Available." *Journal of Economic Perspectives* 26 (4):27-42.

Carson, Richard T., and Theodore Groves. 2007. "Incentive and informational properties of preference questions." *Environmental and Resource Economics* 37:181-210.

Carson, Richard T., and W. Michael Hanemann. 2005. "Contingent Valuation." In *Handbook of Environmental Economics*, edited by Karl-Göran Mäler and Jeffrey R. Vincent, 822-936. Amsterdam: North-Holland.

Carson, Richard T., and Jordan Louviere. 2011. "A Common Nomenclature for Stated Preference Elicitation Approaches." *Environmental & Resource Economics* 49 (539-559).

Charness, Gary, Uri Gneezy, and Michael A. Kuhn. 2012. "Experimental methods: Between-subject and within-subject design." *Journal of Economic Behavior & Organization* 81 (1):1-8. doi: http://dx.doi.org/10.1016/j.jebo.2011.08.009.

Corso, Phaedra S., James K. Hammitt, and John D. Graham. 2001. "Valuing Mortality-Risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation " *Journal of Risk and Uncertainty* 23 (2):165-184. doi: 10.1023/A:1011184119153.

de Blaeij, Arianne, Raymond J.G.M. Florax, Piet Rietvald, and Erik Verhoef. 2003. "The value of statistical life in road safety: a meta-analysis." *Accident Analysis & Prevention* 35 (6):973-986. doi: 10.1016/S0001-4575(02)00105-7

Dekker, Thijs, Roy Brouwer, Marjan Hofkes, and Klaus Moeltner. 2011. "The Effect of Risk Context on the Value of a Statistical Life: a Bayesian Meta-model." *Environmental and Resource Economics* 49 (4):597-624.

Desvousges, William, Kristy Mathews, and Kenneth Train. 2012. "Adequate responsiveness to scope in contingent valuation." *Ecologial Economics* 84:121-128.

Dreze, J.H. 1962. "L'utilité sociale d'une vie humaine." *Revue Francaise de Recherche Opérationnelle* 22.

Fischhoff, Baruch, and Shane Frederick. 1998. "Scope (in)sensitivity in elicited valuations." *Risk, Decision, and Policy* 3:109-123.

Foster, Vivien, and Susana Mourato. 2003. "Elicitation Format and Sensitivity to Scope." *Environmental and Resource Economics* 24 (2):141-160. doi: 10.1023/A:1022856329552.

Frederick, S., G. Loewenstein, and T. O'Donoghue. 2002. "Time discounting and time preference: A critical review." *Journal of Economic Literature* 40 (2):351-401. doi: Doi 10.1257/002205102320161311.

Goldberg, Isabell, and Jutta Roosen. 2007. "Scope insensitivity in health risk reduction studies: A comparison of choice experiments and the contingent valuation method for valuing safer food." *Journal of Risk and Uncertainty* 34:123-144.

Hammitt, James K. 2000. "Evaluating Contingent Valuation on Environmental Health Risks: The Proportionality Test." *AERE Newsletter* 20 (1):14-19.

Hammitt, James K., and John D. Graham. 1999. "Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability?" *Journal of Risk and Uncertainty* 18 (1):33-62. doi: 10.1023/A:1007760327375.

Hammitt, James K., and Kevin Haninger. 2010. "Valuing Fatal Risks to Children and Adults: Effects of Disease, Latency, and Risk Aversion." *Journal of Risk and Uncertainty* 40 (1):57-83.

Haninger, Kevin, and James K. Hammitt. 2011. "Diminishing Willingness to Pay per Quality-Adjusted Life Year: Valuing Acute Foodborne Illness." *Risk Analysis* 31 (9):1363-1380. doi: 10.1111/j.1539-6924.2011.01617.x.

Hanley, Nick, Wiktor Adamowicz, and Robert E. Wright. 2005. "Price vector effects in choice experiments: an empirical test." *Resource and Energy Economics* 27 (3):227-234. doi: http://dx.doi.org/10.1016/j.reseneeco.2004.11.001.

Hattie, John, and Helen Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1):81-112. doi: 10.3102/003465430298487.

Hausman, Jerry A. 2012. "Contingent Valuation: From Dubious to Hopeless." *Journal of Economic Perspectives* 26:43-56.

Hensher, David A., and William Greene. 2003. "A latent class model for discrete choice analysis: contrasts with mixed logit." *Transportation Research Part B* 37:681-698.

Hensher, David A., J.M. Rose, J. deDios Ortuzar, and Luis I. Rizzi. 2009. "Estimating the willingness to pay and value of risk reduction for car occupants in the road environment." *Transportation Research Part A* 43 (7):692-707.

Hultkrantz, Lars, and Mikael Svensson. 2012. "The Value of a Statistical Life in Sweden: A Review of the Empirical Literature." *Health Policy* 108:302-310.

ITU. 2012. *ITU Yearbook of Statistics*. http://www.itu.int/ITU-D/ict/publications/yb/index.html: International Telecommunication Union (United Nations).

Jones-Lee, Michael W. 1974. "The Value of Changes in the Probability of Death or Injury " *Journal of Political Economy,* 82 (4):835-849.

Kahneman, Daniel, and Jack L. Knetsch. 1992. "Valuing public goods: The purchase of moral satisfaction." *Journal of Environmental Economics and Management* 22 (1):57-70. doi: 10.1016/0095-0696(92)90019-S.

Kahneman, Daniel, Ilana Ritov, and David Schkade. 1999. "Economic preferences or Attitude Expressions?: An Analysis of Dollar Responses to Public Issues." *Journal of Risk and Uncertainty* 19 (1-3):203-235. doi: 10.1023/A:1007835629236.

Kalman, Kenneth C., and Geoff Royston. 1997. "Risk language and dialects." *British Medical Journal* 315 (11 October):939-942. doi: PMID: 9361547

Kochi, Ikuho, Bryan Hubbell, and Randall Kramer. 2006. "An Empirical Bayes Approach to Combining and Comparing Estimates of the Value of a Statistical Life for Environmental Policy Analysis." *Environmental & Resource Economics* 34 (3):385-406. doi: 10.1007/s10640-006-9000-8.

Koltowska-Häggström, Maria , Björn Jonsson, Dag Isacson, and Kerstin Bingefor. 2007. "Using EQ-5D To Derive Utilities For The Quality Of Life - Assessment Of Growth Hormone Deciency In Adults (QoL-AGHDA)." *Value in Health* 10 (1):73-81.

Lindhjelm, Henrik, Ståle Navrud, Nils Axel Braathen, and Vincent Biausque. 2011. "Valuing Mortality Risk Reductions from Environmental, Transport, and Health Policies: A Global Meta-Analysis of Stated Preference Studies." *Risk Analysis* 31 (9):1381-1407.

Meyer, A. 2013. "Intertemporal Valuation of River Restoration." *Environmental & Resource Economics* 54 (1):41-61. doi: DOI 10.1007/s10640-012-9580-4.

Mishan, Ezra J. 1971. "Evaluation of Life and Limb: A Theoretical Approach." *Journal of Political Economy* 79 (4):687-705.

NICE. 2011. Discounting of health benefits in special circumstances. National Institute for Health and Clinical Excellence.

Payne, John, James Bettman, David Schkade, Norbert Schwarz, and Robin Gregory. 2000. "Measuring Constructed Preferences: Towards a Building Code." In *Elicitation of Preferences*, edited by Baruch Fischhoff and Charles Manski, 243-275. Springer Netherlands.

Peters, E., N. F. Dieckmann, D. Vastfjall, C. K. Mertz, P. Slovic, and J. H. Hibbard. 2009. "Bringing meaning to numbers: the impact of evaluative categories on decisions." *J Exp Psychol Appl* 15 (3):213-27. doi: 10.1037/a0016978.

Poe, Gregory L., Kelly : Giraud, and John Loomis. 2005. "Computational Methods for Measuring the Difference of Empirical Distributions." *American Journal of Agricultural Economics* 87 (2):353-365.

Quinet, Emile, Luc Baumstark, and et al. 2013. Cost Benefit Assessment of Public Investments: Summary and Recommendations. edited by Commissariat géneral à la stratégie et à la prospective. Paris, France: Policy Planning Commission.

Randall, Alan, and John P. Hoehn. 1996. "Embedding in Market Demand Systems." *Journal of Environmental Economics and Management* 30 (3):369-380.

Rheinberger, Cristoph M. 2011. "A Mixed Logit Approach to Study Preferences for Safety on Alpine Roads." *Environmental and Resource Economics* 49:121-146.

Robinson, Lisa A., and James K. Hammitt. 2011. "Behavioral Economics and the Conduct of Benefit-Cost Analysis: Towards Principles and Standards." *Journal of Benefit-Cost Analysis* 2 (2):Article 5.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differnetiation in Pure Competition." *Journal of Political Economy* 82 (1):34-55.

SCB. 2010. "Statistics Sweden: SCB statistikdatabaser: Hushållens ekonomi (HEK)."

SCB. 2011. "SCB Statistikdatabasen: Befolkningsstatistik (Statistics Sweden: Population Statistics Database)."

Schelling, Thomas C. 1968. "The Life You Save May Be Your Own." In *Problems in Public Expenditure Analysis*, edited by Samuel B. Chase, 127-162. Washington D.C.: The Brookings Institution.

Shogren, Jason F., and Thomas D. Crocker. 1999. "Risk and Its Consequences." *Journal of Environmental Economics and Management* 37 (1):44-51. doi: http://dx.doi.org/10.1006/jeem.1998.1059.

Slothuus Skjoldborg, Ulla, and Dorte Gyrd-Hansen. 2003. "Conjoint analysis. The cost variable: an Achilles' heel?" *Health Economics* 12 (6):479-491.

Slovic, P., J. Monahan, and D. G. MacGregor. 2000. "Violence risk assessment and risk communication: the effects of using actual cases, providing instruction, and employing probability versus frequency formats." *Law Hum Behav* 24 (3):271-96.

Slovic, Paul. 2000. *The Perception of Risk*. London: Earthscan.

Taylor, E.V., K.M C. Herman, E.C. Ailes, C. Fitzgerald, J.S. Yoder, B.E. Mahon, and R.V. Tauze. 2012. "Common source outbreaks of Campylobacter infection in the USA, 1997–2008." *Epidemiology and Infection* forthcoming.

Train, Kenneth. 2009. *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.

Viscusi, Kip, and Jahn K. Hakes. 2004. "Dead Reckoning: Demographic Determinants of the Accuracy of Mortality Risk Perceptions " *Risk Analysis* 24 (3):651-664.

Viscusi, Kip, Joel Huber, and Jason Bell. 2008. "Estimating discount rates for environmental quality from utility-based choice experiments." *Journal of Risk and Uncertainty* 37:199-220.

Viscusi, W. Kip, and Joseph E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World " *Journal of Risk and Uncertainty* 27 (1):5-76. doi: 10.1023/A:1025598106257.

Tables

Table 1 Survey description: attributes and attribute levels

Variable name	**Description**	**Attribute levels**
water	Source of disease	Food = 0, Water = 1
die	Mortality reduction – individuals per year	Sample A: 1, 2, 4 Sample B: 100, 200, 400
sick	Morbidity reduction – individuals per year	8 000, 16 000, 32 000
delay	Delay in years until policy starts to have effect (0 = no delay)	0, 2, 5, 10
cost	Cost in SEK per year	500, 1 000, 2 000

Note: USD 1 = SEK 6.6, 2012-11-12

Table 2 Descriptive Statistics of background variables

Variables	**Description**	**Sub-sample A**	**Sub-sample B**	**Swedish population**
Male	=1 if male	0.50	0.50	0.50
Age	Age in years	45.10	45.22	48.80
		(16.57)	(16.64)	
University Education	=1 if university education ≥ 3 years	0.32	0.34	0.19
Employment	=1 if currently employed (age 18>)	0.58	0.60	0.63
Income	Disposable household income in SEK	18 017	19 483	21 825*
	(USD 1 = SEK 6.6, 2012-11-12)	(8 361)	(9 442)	

Note: Standard deviations in parentheses. Number of respondents in sub-sample A: 1000, and in sub-sample B: 250.
* 2010 median household income.

Table 3 Risk experience and perception: Sub-sample A and B

Variables	Description	Sub-sample A	Sub-sample B
Food poisoned	=1 if food poisoned last year due to any reason	0.08	0.12
Campylobacter	=1 if (ever) food poisoned due to confirmed campylobacter	0.08	0.08
Bottled water	=1 if buys bottled water when in foreign countries	0.69	0.64
Water risk	=1 if subjective good knowledge of water-borne diseases	0.15	0.13
Food risk	=1 if subjective good knowledge of food-borne diseases	0.22	0.22
Public Risk perception	Subjective beliefs regarding annual risk of food poisoning (all causes) (objective average risk 10/100)	17.27/100	16.73/100
		(17.19/100)	(18.28/100)
Individual Risk perception	Subjective beliefs regarding individual risk of campylobacteriosis per year (average objective risk 7/1000).	16.35/1000	25.65/1000
		(81.37/1000)	(118.30/1000)
Health	Health status as measured on a Visual Analog Scale 0-100	80.09	81.94
		(16.77)	(15.44)

Note: Standard deviations in parentheses.

Table 4 Benchmark models

	Sub-sample A	**Sub-sample B**
sq	0.0966	0.00754
	(1.16)	(0.04)
water	0.235^{***}	0.237^{***}
	(6.23)	(3.19)
sick	0.0000309^{***}	0.0000243^{***}
	(13.43)	(5.58)
die	0.298^{***}	0.00344^{***}
	(16.79)	(9.84)
cost	-0.000566^{***}	-0.000442^{***}
	(-15.97)	(-6.56)
delay	0.101^{***}	0.0919^{***}
	(12.31)	(5.86)
Estimated VSL[a]	4 732	70
	(3 954 – 5 509)	(45 – 95)
Estimated VSI[a]	0.49	0.49
	(0.40 – 0.59)	(0.26 - 0.73)
Number of respondents	1003	250
Number of responses	8024	2000
Log-likelihood	-8120.73	-2001.55
AIC	16253	4015
BIC	16283	4036

t statistics adjusted for clustering at the respondent level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

a:In SEK million. 95 % confidence intervals in parentheses.

Table 5 Latent class models

	Sub-sample A		**Sub-sample B**	
	Class 1	**Class 2**	**Class 1**	**Class 2**
sq	-1.929***	1.628***	-2.314***	1.418***
	(-17.05)	(7.75)	(-10.43)	(4.38)
water	0.380***	0.163	0.387***	-0.133
	(6.84)	(1.44)	(3.54)	(-0.62)
sick	0.0000477***	0.0000309***	0.0000353***	0.0000231*
	(17.13)	(5.01)	(6.67)	(1.93)
die	0.396***	0.201***	0.00458***	0.00127
	(20.19)	(4.11)	(11.86)	(1.38)
cost	-0.000585***	-0.00152***	-0.000463***	-0.000980***
	(-19.78)	(-10.81)	(-8.26)	(-4.82)
delay	0.121***	0.122***	0.117***	0.191
	(14.54)	(3.20)	(6.91)	(1.48)
Class probability	0.703***	0.297***	0.735***	0.265***
	(46.51)	(19.67)	(26.22)	(9.43)
Estimated VSL[a]	6 095	1 186	89	12
	(5 314 – 6 876)	(590 – 1 782)	(65 – 113)	(-5 – 29)
Estimated VSI[a]	0.73	0.18	0.69	0.21
	(0.63 – 0.84)	(0.10 – 0.26)	(0.44 - 0.93)	(-0.02 - 0.45)
Number of respondents	1003		250	
Number of responses	8024		2000	
Log-likelihood	-5860.90		-1473.88	
AIC	11748		2974	
BIC	11812		3020	

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

a:In SEK million. 95 % confidence intervals in parentheses.

Figures

Figure 1 Example of Choice Set in sub-sample A

What do you prefer in this situation?

	Policy A	**Policy B**
Source of disease	Water	Food
Number of fewer individuals who die (per year) when the policy is implemented	1	2
Number of fewer individuals who get sick (per year) when the policy is implemented	16 000	8 000
The policy starts to have effect	this year	in 10 years
Your cost (per year)	1 000 SEK	2 000 SEK

I prefer

☐ Policy A

☐ Policy B

☐ None of the suggested policies (today’s situation remains and no additional cost for you)

Note: The choice sets in sub-sample B were identical to the ones in sub-sample A with the exception that the levels of the attribute “fewer individuals who die” were multiplied by 100.

Figure 2 Plot of coefficient ratios in the two models