Sensitivity to Scope in Contingent Valuation and Discrete Choice Experiments: Results Based on Valuing Health Risk Reductions

Abstract

There is a large stated preference literature estimating willingness to pay (WTP) for health risk reductions using the contingent valuation (CVM) approach, and more recently often based on the discrete choice experiment (DCE) approach. Irrespective of method, these studies often fail to show adequate sensitivity to scope, i.e. WTP does not increase as the quantity of the good or the number of goods increases. In this paper we compare the sensitivity to scope with the CVM and DCE approach based on respondents' WTP for mortality and morbidity risk reductions. We analyze scope sensitivity using between-subject tests, which is a novelty in the DCE setting. The results show that we can reject adequate sensitivity to scope in both the CVM and DCE design, and the degrees of bias and welfare estimates are very similar in the two approaches. Thus, using a more stringent scope sensitivity test than the standard approach in the DCE literature indicates that sensitivity to scope is an equal pressing issue in DCE as well as in CVM studies.

Keywords: Discrete choice experiment; Contingent valuation; Morbidity risk; Mortality risk; Scope sensitivity; Willingness to pay

JEL-Codes: D61; H41; I18; Q51

There has been a steady increase in the use of stated preference (SP) methods to estimate willingness to pay (WTP) for non-market goods, both using contingent valuation (CVM) and discrete choice experiments (DCE). In a bibliographic overview Carson (2010) counts to more than 7,500 studies from over 130 countries over the last 50 years. The larger share of SP applications have used CVM designs, although in the last 10-15 years DCE designs have become increasingly common in fields such as environmental, health and transport economics. Some evidence suggests that DCE now is more common to value non-marked goods than the CVM method (Mahieu, Andersson et al. 2014).[1]

Despite the large increase in the applications of SP methods there still exist a substantial degree of controversy regarding its validity and reliability (e.g. Hausman 2012). At center of the criticism are empirical findings suggesting scope insensitivity, i.e. the respondents' willingness to pay (WTP) being insensitive to the size of the good or the number of goods being valued (Fischhoff and Frederick 1998, Murphy, Allen et al. 2005), hypothetical bias, i.e. the tendency to overestimate WTP due to the hypothetical nature of the valuation task (Murphy, Allen et al. 2005), and order and framing effects (Bateman and Langford 1996, Clark and Friesen 2006). Especially insensitivity to scope has been considered a fundamental concern regarding the validity and reliability of SP methods.[2]

---

[1] We assume that readers are familiar with SP methods and the difference between CVM and DCE. For those who are not we recommend Freeman, Herriges et al. (2014), or any other publication on nonmarket valuation or benefit-cost analysis.

[2] By scope insensitivity we here refer to insensitivity of the WTP to the magnitude of a single good, and we let embedding refer to insensitivity of WTP to the number of goods valued (Goldberg and Roosen 2007).

The aim of this paper is to examine sensitivity to scope using both CVM and DCE designs. Given the number of applied SP studies that have been conducted to date there are surprisingly few studies that directly compare CVM and DCE methods in the same valuation context. An early exception were Adamowicz, Boxall et al. (1998) who estimated welfare estimates for caribou preservation using both a DCE and CVM approach. They argued for several merits in the DCE compared to the CVM approach due to, e.g., its capacity to provide a more detailed analysis of respondents' preferences and the smaller variance in the welfare estimates. More recent examples of studies using both CVM and DCE designs and testing for scope sensitivity include Foster and Mourato (2003), who estimated the WTP for charitable services in the UK, and Goldberg and Roosen (2007), who estimated the WTP for safer chicken. The findings in both studies suggested the DCE design to perform better in terms of sensitivity to scope compared with the CVM design. Moreover, Adamowicz, Dupont et al. (2011) examined the value of risk reductions in the context of drinking water quality using CVM and DCE. They reported scope tests according to our definition of embedding (see footnote 3) and found that the WTP for a combination of the two risk reductions examined was higher than for a reduction in one of them for both the DCE and CVM method.

The good of interest in this study is health risk reductions for which we examine scope sensitivity. The problem of insensitivity to scope in SP studies has been argued to be especially severe for health risk reductions (Carson 2012). We add to the literature of examining scope sensitivity in CVM and DCE designs by using a more stringent test for the DCE part, and the primary contribution is that this is the first paper to our knowledge that conducts between-subject scope sensitivity tests in both the CVM and DCE design.

Whereas previously the much more stringent between-subject (external) test has been used for the CVM method, scope sensitivity in DCE studies has been based on within-subject (internal) tests, which may be the reason why empirical evidence has suggested a better sensitivity to scope in DCE compared to CVM studies.

In this study respondents are asked about their preferences for policies to reduce mortality and morbidity risks due to the infectious disease campylobacteriosis. The disease is caused by the bacteria campylobacter and infects humans mainly by contaminated food or water. A number of policies that increase food and water safety have been shown to reduce the associated health risks (Taylor, Herman et al. 2012). The aim of this study is to elicit individual preferences for reducing mortality and morbidity risks due to campylobacteriosis. The specific objectives are: (i) to elicit WTP estimates that can be used for policy purposes, (ii) to examine whether these estimates are robust between elicitation approach (DCE or CVM), and (iii) to examine if DCE and CVM designs show adequate sensitivity to scope when both approaches face the more demanding between-sample scope sensitivity test.[3]

The rest of the paper is organized such that in section 1 we provide a brief description of the concept of WTP to reduce health risks and scope sensitivity tests, together with some empirical findings. The paper proceeds in section 2 with describing the experimental DCE and CVM survey designs, where we highlight that, with the exception of the valuation task either being constructed as a DCE or a CVM task, respondents faced almost identical risk-policy scenarios. The empirical models are

[3] The data used in this paper has partly been analyzed in a recent paper that only focused on the DCE part of the study (Andersson, Hole et al. 2016).

presented in section 3. The results, presented in section 4, show that the welfare estimates are not significantly different in the DCE and CVM design. Regarding the scope sensitivity tests, whereas we find strong evidence of scope sensitivity in the standard DCE-scope test (within-subject test), using the CVM answers we find no scope sensitivity for a morbidity risk reduction and weak evidence of scope sensitivity for a mortality risk reduction. However, when examining scope sensitivity between subsamples (i.e. between-subject tests for both DCE and CVM) the scope sensitivity test fails in the DCE as well as in the CVM design. Thus we cannot conclude that DCE designs are a solution to the frequently reported scope-sensitivity problems in CVM designs. Finally, we also highlight that the typical (within-subject/internal) scope sensitivity test in DCE studies may essentially mask the larger problem of (between-subject) insensitivity to scope.

1. Valuing Safety and Sensitivity to Scope

The aim of the experimental survey in this paper is to elicit the respondents' marginal WTP to reduce mortality and morbidity risk, defined as the value of a statistical life (VSL) and the value of a statistical illness (VSI), which are measures of the population mean marginal rate of substitution between health risk and wealth (e.g.,Dreze 1962, Schelling 1968) and the basis for valuing prevented fatalities and injuries/illnesses in risk-reducing policies and investments. Based on standard assumptions, such that individuals prefer staying alive or being healthy compared to dying or being sick, and that the marginal utility of wealth being higher in a good compared to a bad health state, we expect VSL and VSI to be increasing with wealth, the baseline risk level, and with the size of the risk reduction (Jones-Lee 1974, Weinstein, Shepard et al. 1980, Pratt and Zeckhauser

1996).[4] Moreover, in addition to “standard” scope sensitivity, i.e. that WTP increases with the size of the risk reduction, it can also be shown that the WTP should be near-proportional to the size of the mortality risk reduction. The near-proportionality is thus a necessary, albeit not sufficient, validity criteria for SP based WTP estimates (Corso, Hammitt et al. 2001).[5]

The number of applied SP studies to estimate VSL and VSI is substantial and covers a large range of different risk contexts such as e.g. health risks due to contaminated water (Adamowicz, Dupont et al. 2011, Viscusi, Huber et al. 2012), cancer risks (Hammitt and Haninger 2010), road mortality risks (Andersson, Hammitt et al. 2013) , and fire and drowning risks (Carlsson, Daruvala et al. 2010). In a recent meta-analysis covering only a sub-set of all VSL estimates, Lindhjelm, Navrud et al. (2011) count to almost 1,000 published VSL estimates using SP methods.

The standard and recommended test for scope sensitivity in CVM studies is to let different sub-samples indicate their WTP for different magnitudes of the good (*between-subject scope sensitivity*). The alternative form of scope sensitivity test is when the same respondent is asked to state his/her WTP for different magnitudes of a good in a sequential order (*within-subject scope sensitivity*). A within-subject test is typically seen as a weaker validity test since respondents will show sensitivity to scope if they behave

---

[4] The concept of the VSL (and VSI) and its theoretical model have been presented in several studies. For those not familiar with the concept and its model, see provided references and, e.g., Robinson and Hammitt (2013).

[5] It is important to keep in mind that near-proportionality requires that the baseline and the change in risk are small, and/or that the payment is not a substantial fraction of income. Both of these assumptions are satisfied in our experiment.

internally consistent (Carson, Flores et al. 2001, Bateman, Cole et al. 2004).[6] For this reason the NOAA panel on CVM studies recommended that between-subject, rather than within-subject, scope sensitivity tests should be carried out as a standard procedure to evaluate the validity of estimates of an SP study (Arrow, Solow et al. 1993). For health risk reductions scope sensitivity can also be divided into "weak" and "strong" scope (Corso, Hammitt et al. 2001). The weak scope sensitivity is defined such that WTP should increase with the size of the risk reduction, whereas strong scope sensitivity implies that WTP should increase near-proportionally with the risk reduction.

The empirical literature has documented that several CVM studies fail to pass a between-subject scope sensitivity test (Fischhoff and Frederick 1998, Desvousges, Mathews et al. 2012). Regarding health risks and weak and strong scope sensitivity, earlier research has indicated that strong scale sensitivity is usually not met in CVM studies valuing mortality risk reductions. For instance Hammitt and Graham (1999), found in their review of 25 studies that most studies passed the weak scale sensitivity test, but that none passed the "strong" scope sensitivity test. Whereas some recent studies have now found strong scope sensitivity, the general conclusion in the review has not been challenged (Robinson and Hammitt 2015).

In DCE studies respondents are asked to choose between different alternatives, where the status quo or an opt-out alternative is usually one of the alternatives. The alternatives are described by a group of attributes and the levels of these attributes and respondents are assumed to be more likely to choose alternatives that provide more of

---

[6] An argument in favor of a within-subject scope sensitivity test is that it most closely mimics a real market scenario where individuals indeed may learn to show "internal consistency" by repeated actions.

the attribute. Respondents are usually asked to make several of these choices and it has been argued that DCE has an advantage over CVM since respondents are asked to consider two (or more) nested goods in each choice set, i.e. respondents that are internally consistent will value a larger scope higher (Fischhoff and Frederick 1998, Charness, Gneezy et al. 2012). However, in terms of a scope sensitivity test as a validity criterion, this is more similar to a within-subject scope test in CVM studies, which is, as mentioned, a less stringent test and therefore may be seen as a weaker validity criterion as compared to what e.g. the NOAA panel recommended for CVM studies. Lindhjelm, Navrud et al. (2011) in their review of the VSL literature found that 291 out of 318 estimates passed a within-subject scope test. They also found that only 85 out of 199 estimates from CVM studies passed the scope test in between-subjects tests, but were not able to collect any information on how well DCE studies did in between-subjects scope tests since only CVM studies had been designed to allow for such a test.

2. The Experimental Survey Design

2.1 Survey Administration and Structure

The survey was conducted (administered by the company Scandinfo) in the spring of 2012 on-line using a web-panel of respondents who were randomly recruited to the web-panel by phone (there was no “self-recruitment”) among internet-enabled individuals in Sweden aged 18 and over. Obviously, this does not provide us a random sample of all Swedes, but considering that Sweden has among the highest internet penetration rates in the world (ITU 2012) it may be considered as a country where there are strong arguments for using web-based surveys. In total 1 953 respondents were included, where 1 253 respondents were randomly selected to perform the DCE task and 700 respondents were randomly selected to perform the CVM task.[7]

Respondents were, when invited to take part in the survey, told that their answers would be of relevance to guide society on how to allocate resources. This suggestion of policy relevance of the study was to make sure that the respondents treated the survey as consequential (Carson and Groves 2007). Following standard procedure, the survey was pre-tested in small focus groups as well as using two on-line pilot studies. The reason for using two pilot studies was because the results from the first pilot suggested that some of the attribute levels in the DCE should be modified and to make sure the new levels were adequately set a new pilot study was conducted. Overall the pre-testing led to a number of small textual changes regarding the risk descriptions, and some small modifications to the attribute levels.

The survey consists of four sections that are identical with the exception of the third part which either contains a DCE or a CVM task. Respondents were randomly allocated to either performing a DCE or a CVM task. The first section of the survey contains 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). The second and third section contain the description of the risk scenario and the WTP questions, respectively, and we will describe those two sections in more detail in the following sections of this paper. The fourth section includes questions on

---

[7] It was intended to have 1 000 respondents in one of the samples, but 1 003 respondents answered the survey in that sample which is the reason 1 253 respondents in the DCE subsample.

socio-economics and demographics. Upon completing the four sections the respondents could, if they desired, participate in a number of survey debriefing questions.

2.2 Risk description and scenario

Section two describes the risk of getting campylobacteriosis and its symptoms to the respondents. Respondents are presented with three different severity levels of campylobacteriosis: mild, moderate, and severe. They are provided with the distribution of the different severity levels (i.e. the probability of ending up with one of the severity levels conditional on catching it), together with the symptoms, durations, and whether a physician should be consulted and/or whether hospitalization would be necessary. Since there is no variation in the description of the severity levels and their probabilities, it is not possible to estimate WTP for specific severity levels, or duration of illness, etc.

The different subsamples are described in Table 1. Subsample A consists of DCE_A and CVM_A. It is our main subsample in which we present recent actual baseline mortality and illness risk levels together with possible risk reductions. Subsample B, consisting of DCE_B and CVM_B, has a strict research purpose which is why the number of respondents is smaller. In both A and B, the annual incidence is described to be 63 000 cases in Sweden, which corresponds to a risk of 7 in 1,000 (AgriFood 2012). The risk denominator is the total Swedish population. The description of the mortality baseline risk differs, however, slightly between the samples. In subsample A the respondents are presented with the following information, "In addition, in rare events the illness can lead to death (less than 5 cases in total per year in Sweden)." Sample B is presented with the same phrase except that the text within the parenthesis is left out. The reason is that the number of prevented mortalities in B will be larger than the actual baseline risk, and in order not to provide any incorrect information to the respondents the number 5 was left out in subsample B.

The motivation for leaving out the number for the baseline risk in B is because it allows us to conduct our novel scope sensitivity test and compare our findings to previous findings in the literature. As shown in Table 1 the mortality risk reductions used in subsample A are 1, 2 and 4 fewer deaths per year which were decided after discussions and feedback from a medical expert in the field of infectious diseases, relevance to the research question, and feedback from focus groups. In subsample B these values are multiplied by a factor 100 to make the risk reductions in B in line with the levels in Swedish road VSL studies (Hultkrantz and Svensson 2012).[8] Most studies in Sweden to estimate VSL are based on road safety and by using risk reductions in line with that literature it enables us to compare our VSL estimates with a large range of previously reported Swedish VSL estimates.[9] No change was made between subsamples A and B concerning the risk of illness which is the same in both subsamples.

It is well established that individuals have difficulties understanding small probabilities and to lessen the cognitive burden the risk changes were presented as frequencies instead of probabilities (Kalman and Royston 1997, Slovic, Monahan et al.

---

[8] Note that even if the relative difference between the mortality risk reductions in A and B are large, the absolute differences are not that different from a layperson's perspective. Hence, we do not expect the difference in the size of risk reductions to cause any problems in our design, and 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.

[9] To clarify, there was no mentioning of road-fatality risk in sample B.

2000). Moreover, with the Swedish population as the risk denominator the baseline risks are presented as verbal probability analogs, but for the morbidity risks we also described the number of incidences as the number of cases in an average sized city of Sweden, i.e. 700 cases in a city of 100,000. Due to the very small baseline mortality risk in subsample B, that no precise number was presented in subsample B, and that the evidence of visual aids' capability to address scope insensitivity is mixed (Corso, Hammitt et al. 2001, Goldberg and Roosen 2007, Haninger and Hammitt 2011), we decided not to use any visual aid in the survey. In the DCE subsamples respondents were also provided feedback after their first choice set as an instrument to increase their understanding of the risk-money dollar tradeoffs they faced, which is explained in section 2.3.1.

[Table 1 about here]

2.3 WTP section

Respondents are asked about their preferences for different public policies that are described to reduce campylobacter-related mortality and morbidity risks and either targeting food- or water-borne risk. The policy scenario is identical between the DCE and the CVM with respect to the size of mortality and morbidity risk reductions, and the source of the disease being targeted (food- or water-borne). Individuals who only care about safety, no matter how it can be achieved, would not care about whether food or water is targeted. However, respondents' preferences may be influenced if they perceive the risks to differ with respect to controllability, dreadfulness, etc. (Shogren and Crocker

1999, Slovic 2000). We assume that water risk may be perceived as less controllable which is negative from the individual's perspective

A difference between the DCE and CVM designs is that in the DCE task respondents are also asked to take into account when the policy will start to provide benefits when choosing between the different alternatives. This dimension was not included in the CVM sample due to a restriction on how many respondents that could be included in the survey. Regarding the payment mechanism it may seem reasonable to have different mechanisms for food and water safety, but to avoid having confounding effects from different payment mechanism on the interpretation of our results we decided to use a generic payment mechanism. The cost and risk attributes are presented as the annual cost and the annual reduction in cases to make it easier for respondents to interpret and understand them, whereas the policies are presented to have a 5-year duration to increase the realism of the scenario.[10]

2.3.1 The DCE Task

In the DCE, the policies differ 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. The attributes and their levels are shown in Table 2. The delay attribute is included to elicit individual time preferences. The delay only concerns when the policy will start to have an effect, i.e. there is no delay in the payment of the cost of the policy.

---

[10] Respondents were informed that the social insurance system would compensate potential income losses and health care costs (if becoming sick) to make sure that elicited preferences reflect the health and cost domains as presented in the scenarios.

As explained, the mortality risk reductions were chosen to be reasonable based on the actual baseline risk in DCE_A and to be in line with the size used for road-mortality risk reductions in earlier Swedish WTP studies in DCE_B. The levels for the morbidity risk reductions were then 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 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 VSI (Lindhjelm, Navrud et al. 2011). The levels for the risk and cost attributes were discussed in the focus groups 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) allowing for all possible two-way interactions to be estimated. The 64 choice sets were randomly blocked into eight versions, which imply that each respondent faced eight choice sets.

[Table 2 about here]

Before the respondents face the choice sets a general description of the policy scenario is included, stated as (freely translated from Swedish):

> *"Assume 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".

An example of a choice set, as faced by respondents in DCE_A, is shown in Figure 1. As shown, the respondents are 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. After the respondent's first choice he/she is provided some feedback on the computer screen on the meaning of his/her choice regarding changes in risk, the cost associated, etc. The respondent is then asked if he/she is happy with his/her choice and want to proceed to the next choice set or change the choice in the current choice set. We found that 16.8 percent of the respondents changed their initial choice. In the following 7 choice sets respondents are not given the possibility to change their decisions. Moreover, no "reading ahead" is possible, i.e. respondents have to answer the current choice set before moving on to the next one.

[Figure 1 about here]

Scope sensitivity is analyzed using both within-subject and between-subject test. The former is the standard scope sensitivity test in the DCE literature, whereas the latter is the novelty of our experiment, and is conducted by comparing welfare estimates between samples DCE_A and DCE_B. This comparison is not part of the typical validity test in the DCE literature but is actually in line with classic validity tests dating back to e.g. the NOAA panel (Arrow, Solow et al. 1993).

2.3.2 The CVM Task

The scenarios in CVM_A and CVM_B are identical to those in the DCE questionnaires expect that no delay aspect of when the policy would take effect is included. The split sample design on mortality risks mimics then one in the DCE, and thus the same levels for the risk reductions are used in the two subsamples, i.e. in CVM_A smaller reductions (1, 2, or 4 fewer deaths), and in CVM_B larger reductions (100, 200, or 400). Since only one attribute can vary at the time in a CVM we ask each respondent four WTP questions[11]: (1) food mortality reduction, (2) food morbidity reduction, (3) water mortality reduction, (4) water morbidity reduction. This enables us to conduct both within- and between-subjects test as in the DCE.[12] The different subsamples and the order of the questions are shown in Table 3.

We choose to ask the respondents an open-ended WTP question in the CVM. The motivation for choosing the open-ended format is to: (i) avoid anchoring effects (Green, Jacowitz et al. 1998), and (ii) to contrast the DCE, where the analysts set all the levels, with an open-ended CVM where respondents have total freedom to state any number. The latter reason makes sure that we do not influence the comparison of our methods by choosing the bid levels in the CVM.

[Table 3 about here]

Before the WTP question respondents face a general description of the policy scenario which is formulated to be as identical to the DCE scenario as possible (freely translated from Swedish):

---

[11] For the purpose of identification.
[12] Anchoring and starting point bias are well known issues in the CVM. Our design allows us to conduct within-subject tests, but as our results suggests we may have issues with our second to fourth answer, and in our analysis we only use data from respondents' first answer.

"Assume 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 4 different questions".

Hence, the difference compared to the DCE scenario is confined to excluding the delay attribute, that respondents' WTP for the type of risk reduction (water or food and mortality and morbidity) is elicited in separate questions, and that instead of using a cost attribute the respondents' WTP is elicited with a direct question. Using mortality risk as the example the WTP question to reduce mortality risk is then posed as:

"*What is the maximum amount that you would be willing to pay per year during a five year period for a stricter [food, water] sanitation that would imply that [1, 2, 4, 100, 200, 400] fewer persons per year will die due to campylobacteriosis?"*

For the scope sensitivity tests with our CVM data we used both descriptive statistics as well as regression analysis.

3. Empirical models

For the DCE valuation task we provide a brief specification of the empirical model. As described above, respondents are asked to choose their preferred option between two hypothetical scenarios and the status-quo, i.e. a total of *J*=3 alternatives, in *T*=8 choice sets. The utility that respondent *n* derives from choosing alternative *j* in choice set *t* in our specification is given by

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

where $\beta_0, \dots, \beta_5$ are coefficients to be estimated, sq_{njt} is an alternative-specific constant for the status quo alternative, *die, sick, delay, water*, and *cost* the attributes as described in Table 2, and ε_{njt} is a random error term which is assumed to be IID type I extreme value. To examine the effect on individual characteristics on the WTP we also run regressions with age and income groups interacted with the attributes. The specification of the delay attributes means that δ is the discount rate for mortality and morbidity (Alberini and Ščasný 2011).

As explained, the marginal WTP is given by the marginal rate of substitution between wealth and risk, which can be estimated using the regression results from our model. The VSL, which can be interpreted as the WTP for a reduction in risk equivalent to saving one life, is given by

$$-\frac{\partial U_{njt} / \partial die_{njt}}{\partial U_{njt} / \partial cost_{njt}} = -\frac{\beta_1}{\beta_4}. \quad (2)$$

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.[13]

In the CVM design we have access to precise estimates of respondents' stated WTP and we therefore directly can estimate the sample mean WTP, VSL and VSI. Let the value of a statistical case (VSC) denote both VSL and VSI, then VSC is estimated as the ratio between the maximum WTP and the change in risk,

---

[13] Andersson, Hole et al. (2016) also examined preference heterogeneity and its potential effect on scope sensitivity by in addition to the conditional logit also run the mixed logit and latent class models. Whereas the results for the former were not robust enough to draw any conclusion, results for the latter provided the same general interpretation of the findings as the conditional logit estimates that we refer to the later in this paper.

$$\boldsymbol{VSC} = \frac{\boldsymbol{WTP}}{\Delta \boldsymbol{p}}, \quad (3)$$

which is then used to estimate the population mean. Moreover, since we have continuous data on the respondent's WTP we can estimate an OLS WTP regression to examine how different covariates influence WTP,

$$\ln(\boldsymbol{WTP_s}) = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} \ln(\Delta \boldsymbol{p_s}) + \sum_{\boldsymbol{k}} \boldsymbol{\beta_k x_k} + \boldsymbol{\varepsilon}. \quad (4)$$

The regressions are run as a log-linear model, which allows us to test whether the WTP is proportional to the size of the risk reduction (Δp), i.e. $\beta_1=1$. We run the WTP regressions separately for mortality and the morbidity risk reduction, s = (mortality, morbidity), and we run regressions with the same individual characteristics as in the conditional logit on the DCE data.

4. Results

4.1 Descriptive Statistics

Table 4 shows descriptive statistics for sex, age, university education, employment, income, and self-reported health status for subsamples DCE_A, DCE_B, CVM_A, and CVM_B. There are about as many men as women in the sample and the mean age is around 45. Comparing the different sub-samples, we find that there are no statistically significant differences and the different groups are balanced with respect to the demographic observable characteristics. Our sample corresponds well to national statistics, with the exception of the share of individuals with a university education (3 years or more); with 32-37 percent of our sample having a university education compared to 19 percent in the Swedish population (in the age range 18+). Regarding self-reported health status the estimates are in line with previous Swedish studies (see Andersson, Hole et al. 2016 for references).

Table 4 also includes descriptive statistics on four risk variables collected in the survey. Again, levels are very close between subsamples and there is no statistical significant difference in any of these variables between our subsamples. Moreover, public and individual risk perception is quite good, with an accurate perception of the public and a slight underestimation of the average individual risk. This could be (speculatively) interpreted as the respondents having a good understanding of the risks and/or having taken the survey seriously.

[Table 4 about here]

4.2 DCE Results

In Table 5 we show the regression results from the conditional logit estimation on samples DCE_A and DCE_B. We see that respondents in both samples prefer policies with lower cost and larger mortality and morbidity reductions. This implies that in both samples respondents show within-subject scope sensitivity in line with the theoretical predictions.

Regarding other results we find that respondents have a preference for policies that are water rather than food-based, which is in line with the hypothesis that water risk is perceived less controllable than food risk, and for policies that come into effect sooner rather than later. The result for *delay* suggests a discount rate between 9-10%, except for Reg DCE 4 where it is 3%. Findings from recent studies suggest a discount rate in the rage 7-14% (Alberini, Tonin et al. 2007, Viscusi, Huber et al. 2008, Rheinberger

2011, Meyer 2013) which is in line with the results from our study. The typical respondent also has no preference for or against the status-quo alternative.[14] Hence, the findings from all models are in line with expectations, which we could argue suggest valid estimates of people's preferences.

[Table 5 about here]

The models with interactions for age and income (reference groups are age group 34 years old and younger and with an income below SEK 30000) show that only few of the interactions are statistically significant. We find that the oldest age group has a higher discount rate, but apart from that there are no trends that can be observed from the regression results.

4.3 CVM Results

The first step of the analysis of the WTP data from the CVM study was to examine the scope sensitivity for each type of risk, independent of the order in which the respondents answered the WTP questions. For instance, using the WTP for food-mortality risk for all respondents, whether question asked first, second, third, or forth (see Table 2), do the results suggest that the WTP is scope sensitive? The evidence suggests no scope sensitivity, either within or between subsamples (CVM_A and CVM_B) for any of our risk types. Based on this finding, together with previous evidence in the literature on anchoring and starting point bias, we decided to only use the respondents' first WTP answer.

---

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

Table 6 shows the mean estimate for WTP according to types of risk, food or water and mortality or morbidity, and the risk reduction levels. As can be seen, the mean WTP does not vary to a very large degree across the risk reductions, giving an indication of failure to satisfy a scope sensitivity test. Holding type of risk constant and running a scope sensitivity test for each subsample (DCE_A and DCE_B, and subsamples 1 to 4 in Table 2) we find no evidence of any scope sensitivity.

[Table 6 about here]

The results from the regression analyses are shown in Table 7. The results suggest that respondents have strong preferences for food safety relatively to water safety, i.e. the coefficient on *source* is negative and highly statistically significant in all regressions. This contradicts the findings in the DCE models where preferences for water safety were stronger. The regressions with individual characteristics (*Reg CVM 2* and *4*) use the same individual variables as the conditional logit model for the DCE data. The findings suggest that richer individuals have a higher WTP, but that age has no effect. Regarding scope sensitivity, we only find a positive and statistically significant relationship in one regression, i.e. *Reg CVM 2*. The results in this regression suggest that respondents' WTP increases with the size of the risk reduction, but that it is not proportional to the size. In the same model we also find a negative and statistically significant coefficient for CVM_B which is a dummy for observations from that subsample and included to examine the effect from pooling the observations from the different subsamples.

[Table 7 about here]

4.4 Value of a statistical life and illness estimates

We report the VSL and VSI estimates from the DCE and CVM designs in Table 8. The top half of the table contains the values from the DCE design, whereas the bottom half contains the results from the CVM design.

First focusing on the DCE estimates the importance of a thorough and critical examination becomes apparent. Whereas the regression findings in Table 5 suggest valid estimates of individual preferences (correct and statistically significant signs), the VSL estimates suggest no or very low scope sensitivity between the two subsamples. According to the model estimated for DCE_A the VSL is SEK 4 732 million (USD 710 million), while according to the model estimated for DCE_B the VSL is SEK 70 million (USD 11 million).[15] Hence, with a 100 times smaller risk reduction in DCE_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 et al. (2005) we can reject the null hypothesis that the VSL in DCE_A 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.[16] 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.

---

[15] Estimates based on models without interactions, i.e. *Reg DCE 1* and *3*.

[16] 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.

The results for the CVM subsamples are similar as for the DCE. The first observation is that the VSI estimates are robust between the subsamples, as was the case in the DCE. The second observation is that VSL estimates are in line with the findings from the DCE. Whereas the estimate in CVM_B is in the range of previous Swedish estimates from the road safety literature, SEK 34 million (USD 5 million) the estimate from CVM_A is considerably higher, SEK 3 344 million (USD 507 million). Moreover, we now find that with a 100 times smaller risk reduction we do get a VSL that is 100 times larger. That is, the WTP is very similar between the subsamples.

[Table 8 about here]

5. Conclusion

This paper used a novel approach to conduct a between-subject scope sensitivity test using DCE and CVM designs in valuation of mortality and morbidity risk reductions. Using the standard scope sensitivity test in the DCE literature (within-subject) we document a strong and statistically significant sensitivity to scope. However, in our between-subject analysis we do not find a satisfactory sensitivity to scope. The welfare estimate of VSL is 4 372 million SEK based on the smaller risk reduction and 70 million SEK based on the larger risk reduction, indicating a severe scope insensitivity. In the CVM design we also fail to find satisfactorily scope sensitivity in the identical between-subject analysis, with VSL estimated to be 3 344 million SEK (small risk reduction) and 34 million SEK (large risk reduction). Hence, using a between-subject scope sensitivity test of magnitude, we fail to find proper scope sensitivity in both the DCE and CVM approach, even though we find significant scope effects in the typical within-subject scope test as carried out in the DCE literature today.

The results suggest that the typical test of scope insensitivity in DCEs may not be sufficient to really examine scope sensitivity. As an example, Goldberg and Rosen (2007) tested sensitivity to scope in CVM and DCE in a context of salmonellosis and campylobacteriosis based on a small German sample. They found weak sensitivity to scope, i.e. larger WTP for larger risk reductions (but not proportionally larger), in the DCE approach (but not CVM). This was based on the finding that the risk reduction attribute was statistically significant and with the expected sign, i.e. respondents in the sample preferred policies with a larger risk reduction compared to a smaller one. This is exactly what we find in our study as well, i.e. we can reject within-sample weak scope insensitivity in DCE_A and DCE_B. However, in the comparison between DCE_A and DCE_B we identify a substantial scope insensitivity problem.

In sum, despite the fact that the DCE finds evidence of strong support for weak scope sensitivity using standard tests, both methods perform equally poorly in a between-subject test in our novel between sample tests. Hence, we have highlighted that even though a DCE study finds scope sensitivity based on given choice sets, this does not necessarily mean that valid estimates of individual WTP/preferences have been elicited. The problem often documented in CVM studies with insensitivity to scope is perhaps just as severe in the DCE literature; it has just not been tested using the same strict between-subject test. As discussed by 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 et al. 2003), precisely what we find in our study.

The estimates of VSL based on DCE_A and CVM_A, i.e. the subsamples were the mortality risk scenario (baseline and risk reductions) is accurate, are considerably higher than previous estimates in the literature. However, the estimates of VSL in DCE_B and CVM_B are both within the USD 2.1-13.7 million (2013 price level) range based on international evidence from SP and revealed preference studies (Robinson and Hammitt 2015) and the USD 0.9-10.6 million (2010 price level) range from Swedish road-safety studies (Hultkrantz and Svensson 2012). The fact that we obtain VSL estimates for reducing risk of cambylobacteriosis related to food and water in subsample B (where the risks were designed based on levels used in Swedish WTP road safety studies) in line with VSL values for road safety in Sweden, together with not being able to reject an equal mean of WTP between subsample A and B with the CVM data, is interesting from a reliability perspective. That is, can an estimate of VSL that fall within the range of previous VSL estimates be interpreted as a reliable estimate of individual preferences, or does it only suggest that with the same risk scenario (same levels of risk reductions) similar findings will appear. That is, respondents have an amount to spare for safety projects, no matter the size of the risk reduction? The results from the CVM in this study may suggest that is the case, but it need to be further examined before any conclusions can be drawn.

We agree with Bateman and Brouwer (2006) that “[I]n testing the scope sensitivity of WTP too much emphasis has been placed upon the ‘mere’ demonstration of statistically significant changes in values as levels of provision alter. While such tests are clearly necessary they are far from sufficient.”, and this study has pointed out a potential weakness of how DCE methods are implemented today when eliciting preferences for risk reductions. The study has also shown that the issue is as severe in the CVM (as implemented here). However, this study do not provide any evidence of which method, DCE or CVM, is most useful to elicit individuals’ preferences for small risk reductions, and we do not provide any recommendations. We do believe that the findings in this study can encourage work in that direction, and hence DCE and CVM designs to examine whether one of the methods is systematically better at eliciting preferences for small risk reductions would be a potential direction for future work in this area.

References

Adamowicz, W., et al. (1998). "Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation." American Journal of Agricultural Economics **80**(1): 64-75.

Adamowicz, W., et al. (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. and M. Šcasný (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.

Alberini, A., et al. (2007). "Paying for permanence: Public preferences for contaminated site cleanup." Journal of Risk and Uncertainty **34**(2): 155-178.

Andersson, H., et al. (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, H., et al. (2016). "Valuation of small and multiple health risks: A critical analysis of SP data applied to food and water safety." Journal of Environmental Economics and Management **75**: 41-53.

Ariely, D., et al. (2003). "Coherent Arbitrariness: Stable Demand Curves without Stable Preferences." Quarterly Journal of Economics **118**(1): 73-105.

Arrow, K. J., et al. (1993). "Report of the NOAA Panel on Contingent Valuation." 58 Federal Register 46.

Bateman, I. and R. Brouwer (2006). "Consistency and construction in stated WTP for health risk reductions: A novel scope-sensitivity test." Resource and Energy Economics **28**: 199-214.

Bateman, I., et al. (2004). "On visible choice sets and scope sensitivity." Journal of Environmental Economics and Management **47**: 71-93.

Bateman, I. and I. H. Langford (1996). "Budget-constraint, temporal, and question-ordering effects in contingent valuation studies." Environmental Planning A **28**: 1215-1228.

Bech, M. and D. Gyrd-Hansen (2005). "Effects coding in discrete choice experiments." Health Economics **14**: 1079-1083.

Carlsson, F., et al. (2010). "Value of statistical life and cause of accident: A choice experiment." Risk Analysis **30**(6): 975-986.

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

Carson, R. T. (2010). Contingent Valuation: a comprehensive bibliography and history. Northampton, Edward Elgar.

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

Carson, R. T., et al. (2001). "Contingent Valuation: Controversies and Evidence " Environmental and Resource Economics **19**(2): 173-210.

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

Charness, G., et al. (2012). "Experimental methods: between-subject and within subject design." Journal of Economic Behavior & Organization **81**(1): 1-8.

Clark, J. E. and L. Friesen (2006). The Causes of Order Effects in Contingent Valuation Surveys: An Experimental Investigation, Dept. of Economics, University of Canterbury, New Zealand.

Corso, P. S., et al. (2001). "Valuing Mortality-Risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation " Journal of Risk and Uncertainty **23**(2): 165-184.

Desvousges, W., et al. (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, B. and S. Frederick (1998). "Scope (in)sensitivity in elicited valuations." Risk, Decision, and Policy **3**: 109-123.

Foster, V. and S. Mourato (2003). "Elicitation Format and Sensitivity to Scope." Environmental and Resource Economics **24**(2): 141-160.

Freeman, A. M., et al. (2014). The Measurement of Environmental and Resource Values. New York, NY, USA, RFF Press.

Goldberg, I. and J. 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.

Green, D., et al. (1998). "Referendum contingent valuation anchoring and willingness to pay for public goods." Resource and Energy Economics **20**(2): 85-116.

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

Hammitt, J. K. and K. 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, K. and J. K. Hammitt (2011). "Diminishing willingness to pay per quality-adjusted life years: valuing acute foodborne illness." Risk Analysis **31**(9): 1363-1380.

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

Hultkrantz, L. and M. Svensson (2012). "The Value of a Statistical Life in Sweden: A review of the empirical literature." Health Policy **108**: 302-310.

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

Kalman, K. C. and G. Royston (1997). "Risk language and dialects." British Medical Journal **315**(11 October): 939-942.

Lindhjelm, H., et al. (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.

Mahieu, P.-A., et al. (2014). Is choice experiment becoming more popular than contingent valuation? A systematic review in agriculture, environment and health FAERE Working papers.

Meyer, A. (2013). "Intertemporal Valuation of River Restoration." Environmental & Resource Economics **54**(1): 41-61.

Murphy, J., et al. (2005). "A Meta-analysis of Hypothetical Bias in Stated Preference Valuation " Environmental and Resource Economics **30**(3): 313-325.

Poe, G. L., et al. (2005). "Computational Methods for Measuring the Difference of Empirical Distributions." American Journal of Agricultural Economics **87**(2): 353-365.

Pratt, J. W. and R. J. Zeckhauser (1996). "Willingness to Pay and the Distribution of Risk and Wealth " Journal of Political Economy **104**(4): 747-763.

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

Robinson, L. A. and J. K. Hammitt (2013). "Skills of the trade: valuing health risk reductions in benefit-cost analysis." Journal of Benefit-Cost Analysis **4**(1): 107-130.

Robinson, L. A. and J. K. Hammitt (2015). "Research Synthesis and the Value per Statistical Life." Risk Analysis **35**(6): 1086-1100.

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

SCB (2012). "Statistics Sweden (SCB): Longitudinal integration database for health insurance and labour market studies (LISA)."

Schelling, T. C. (1968). The Life You Save May Be Your Own. Problems in Public Expenditure Analysis. S. B. Chase. Washington D.C., The Brookings Institution: 127-162.

Shogren, J. F. and T. D. Crocker (1999). "Risk and its consequences." Journal of Environmental Economics and Management **37**(1): 44-51.

Slovic, P. (2000). The Perception of Risk. London, Earthscan.

Slovic, P., et al. (2000). "Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats." Law and Human Behavior **24**(3): 271-296.

Taylor, E. V., et al. (2012). "Common source outbreaks of Campylobacter infection in the USA, 1997–2008." Epidemiology and Infection **141**(5): 987-996.

Viscusi, K., et al. (2008). "Estimating discount rates for environmental quality from utility-based choice experiments." Journal of Risk and Uncertainty **37**: 199-220.

Viscusi, W. K., et al. (2012). "Heterogeneity in Values of Morbidity Risks from Drinking Water." Environmental & Resource Economics **52**: 23-48.

Weinstein, M. C., et al. (1980). "The Economic Value of Changing Mortality Probabilities: A Decision-Theoretic Approach." Quarterly Journal of Economics **94**(2): 373-396.

Table 1 Subsample sizes and risk reductions

	Subsample	
Variable name	**A**	**B**
Mortality reduction	1, 2, 4	100, 200 ,400
Morbidity reduction	8 000, 16 000, 32 000	8 000, 16 000, 32 000
N	DCE_A = 1003	DCE_B = 250
	CVM_A = 500	CVM_B = 200

Note: Mortality and morbidity reductions refer to number of individuals per year

Table 2 Attributes and attribute levels in the DCE design

Variable name	Description	Attribute levels
water	Source of disease	Food = 0, Water = 1
die	Mortality reduction – individuals per year	Sample DCE_A: 1, 2, 4
		Sample DCE_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 3 Question order of subsamples in CVM survey

	Food		Water	
Subsample	**Mortality**	**Illness**	**Mortality**	**Illness**
1	1	2	3	4
2	3	4	1	2
3	2	1	4	3
4	4	3	2	1

Table 4 Descriptive statistics in the DCE- and CVM-sample

		DCE Study		CV Study	
Variables	**Description**	**DCE_A**	**DCE_B**	**CVM_A**	**CVM_B**
Demographic and health variables					
Male	=1 if man [0.50]	0.50	0.50	0.49	0.49
Age	Age in years [48.80]	45.04 (16.48)	45.22 (16.64)	45.27 (16.90)	46.17 (17.19)
University	=1 if university ≥ 3 year [0.19]	0.32	0.34	0.30	0.37
Employment	=1 if employed [0.63]	0.57	0.59	0.56	0.58
Income	Disposable household income in SEK per month [33 914]	32 482 (14,878)	34 483 (14,883)	32 962 (14,483)	32 895 (15,444)
Health	Health status as measured on a Visual Analog Scale 0-100	80.07 (16.77)	81.94 (15.44)	80.18 (16.93)	81.28 (14.49)
Risk variables					
Food poisoned	=1 if food poisoned last year due to any reason	0.08	0.11	0.07	0.11
Cambylobacter	=1 if (ever) food poisoned due to confirmed campylobacter	0.07	0.06	0.07	0.08
Public risk*	Subjective beliefs regarding annual risk of food poisoning (all causes) (objective average risk 10/100)	10.40	9.88	9.74	9.78
Individual risk*	Subjective beliefs regarding individual risk of campylobacteriosis per year (average objective risk 7/1000).	3.85	4.03	3.80	4.38
N		1001	250	500	198

Notes: Standard deviations in parentheses. * Geometric mean, and hence N differ when zero answers. No significant differences between subsamples for any of the variables (Kruskal –Wallis). General population statistics in squared brackets (SCB 2011, SCB 2012).

Table 5 Conditinal logit estimates sample DCE_A and DCE_B

	DCE_A		DCE_B	
	Reg DCE 1	**Reg DCE 2**	**Reg DCE 3**	**Reg DCE 4**
sq_asc	0.0966	-0.145	0.00754	0.619*
	(1.16)	(-0.87)	(0.04)	(1.78)
water	0.235***	0.145*	0.237***	0.135
	(6.23)	(1.92)	(3.19)	(0.96)
sick	0.0000309***	0.0000329***	0.0000243***	0.0000262***
	(13.43)	(7.23)	(5.58)	(2.83)
die	0.298***	0.275***	0.00344***	0.00482***
	(16.79)	(7.33)	(9.84)	(6.64)
cost	-0.000566***	-0.000696***	-0.000442***	-0.000597***
	(-15.97)	(-10.29)	(-6.56)	(-4.02)
delay	0.101***	0.0815***	0.0919***	0.0315*
	(12.31)	(6.29)	(5.86)	(1.90)
sq_asc × age 35-54		0.332		-0.789*
		(1.58)		(-1.79)
source × age 35-54		0.0909		0.0938
		(0.97)		(0.50)
sick × age 35-54		0.000000141		-0.0000111
		(0.03)		(-1.05)
die × age 35-54		0.0275		-0.00115
		(0.59)		(-1.23)
cost × age 35-54		0.000117		0.0000542
		(1.36)		(0.32)
delay × age 35-54		0.0250		0.0380
		(1.27)		(1.15)
sq_asc × age 55+		0.0998		-1.404***
		(0.47)		(-3.25)
source × age 55+		0.201**		0.00619
		(2.07)		(0.03)
sick × age 55+		-0.00000854		-0.00000137
		(-1.40)		(-0.12)
die × age 55+		-0.0284		-0.00154*
		(-0.62)		(-1.80)
cost × age 55+		0.0000851		-0.00000348
		(0.93)		(-0.02)
delay × age 55+		0.0537**		0.148**
		(2.11)		(2.44)
sq_asc × SEK 30000+		0.127		0.145
		(0.72)		(0.36)
source × SEK 30000+		0.00620		0.106
		(0.08)		(0.65)
sick × SEK 30000+		0.000000606		0.00000698
		(0.13)		(0.65)
die × SEK 30000+		0.0427		-0.000461
		(1.12)		(-0.58)
cost × SEK 30000+		0.0000884		0.000204
		(1.18)		(1.39)
delay × SEK 30000+		-0.00382		0.0388
		(-0.20)		(1.20)
N (respondents)	1003	986	250	247
N (responses)	8024	7888	2000	1976
Log-likelihood	−8120.73	-7948.64	−2001.55	-1945.10

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6 Descriptive statistics of WTP in CVM

	Food		Water		
Δp	**WTP**	**S.D.**	**WTP**	**S.D.**	**N**
Mortality					
1	832.52	2190.44	650.31	1530.52	167
2	671.13	1441.84	684.54	1513.95	166
4	882.48	1794.42	805.26	1515.23	162
100	897.03	1707.05	1044.88	1888.75	67
200	889.42	2097.55	679.42	1569.70	66
400	483.67	741.21	532.87	793.80	63
Illness					
8000	832.31	1723.38	806.41	1623.87	234
16000	796.13	1783.57	729.70	1522.65	232
32000	895.97	1769.59	803.51	1515.13	225

Note: WTP is in SEK and refers to the arithmetic mean of the unconstrained sample.

Table 7 Regression results CVM

	Mortality risk		Illness risk	
	Reg CVM 1	**Reg CVM 2**	**Reg CVM 3**	**Reg CVM 4**
Δp	0.248	0.379**	-0.203	-0.105
	(0.164)	(0.156)	(0.136)	(0.137)
water	-0.981***	-1.029***	-0.482***	-0.473***
	(0.178)	(0.173)	(0.162)	(0.161)
CVM_B	-0.945	-1.624**	-0.257	-0.222
	(0.784)	(0.737)	(0.185)	(0.181)
SEK 30000+		0.519***		0.603***
		(0.170)		(0.172)
age 35-54		0.138		0.002
		(0.205)		(0.201)
age 55+		0.271		0.319
		(0.211)		(0.197)
Intercept	5.922***	5.418***	8.145***	6.696***
	(0.178)	(0.220)	(1.336)	(1.351)
N	289	285	299	300
Log-likelihood	-526.57	-504.22	-521.65	-520.02
Adjusted R^2	0.095	0.132	0.031	0.071

Standard errors in parenthesis
note: *** p<0.01, ** p<0.05, * p<0.1

Table 8 Value of statistical case estimates in million SEK

	Source	VSC	[95% C.I.]		N
			DCE		
DCE_A					1003
Mortality	Food	4,732	3,954	5,509	
Illness	Food	0.49	0.40	0.59	
DCE_B					250
Mortality	Food	70	45	95	
Illness	Food	0.49	0.26	0.73	
			CVM[a]		
CVM_A					
Mortality	Food	3,344	2,457	4,231	117
Illness	Food	0.567	0.364	0.770	122
CVM_B					
Mortality	Food	34.27	11.18	57.36	47
Illness	Food	0.673	0.194	1.151	48

VSC in million SEK, USD 1 = SEK 6.6 (2012-11-12)
VSC(Mortality)=VSL and VSC(Illness)=VSI
a: Restricted sample: Below 5th and above 95th percentile excluded. Since the DCE specification does not provide a unique VSC for water safety we also report only food VSC from the CVM.

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.