Monograph:
Andersson, H., Hole, A.R. and Svensson, M. (2014) Valuation of small and multiple health risks: a critical analysis of SP data applied to food and water safety. Working Paper. Sheffield Economics Research Paper (201400). Department of Economics, University of Sheffield ISSN 1749-8368

The University Of Sheffield.

Department Of Economics.

Sheffield Economic Research Paper Series.

Valuation of Small and Multiple Health Risks: A Critical Analysis of SP Data Applied to Food and Water Safety

Henrik Andersson
Arne Risa Hole
Mikael Svensson

ISSN 1749-8368

SERPS no. 2014005
April 2014

Inst EAD. Institute for Economic Analysis of Decision-making.

www.sheffield.ac.uk/economics instead.group.shef.ac.uk

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

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

Arne Risa Hole, University of Sheffield

Mikael Svensson, Karlstad University

Abstract: This study elicits individual preferences for reducing morbidity and mortality risk in the context of an infectious disease (campylobacter) using choice experiments. Respondents are in the survey asked to choose between different policies that, in addition to the two health risks, also vary with respect to source of disease being targeted (food or water), when the policy takes place (in time), and the monetary cost. Our results in our baseline model 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. However, a main objective of this study is to examine scope sensitivity of mortality risk reductions using a novel approach. Our results from a split-sample design suggest that the value of the mortality risk reduction, defined as the value of a statistical life, is SEK 3 177 (USD 483 million) and SEK 50 million (USD 8 million), respectively, in our two sub-samples. This result cast doubt on the standard scope sensitivity tests in choice experiments, and the results also cast doubt on the validity and reliability of VSL estimates based on stated preference (and revealed 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, such as benefit-cost analysis.

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

JEL-Codes: D61; H41; I18 ; Q51

* Corresponding author: Toulouse School of Economics (LERNA), 21 Allée de Brienne, 31015 Toulouse Cedex 6, France, Email: henrik.andersson@tse-fr.eu

1. Introduction

The use of benefit-cost analysis (BCA) to guide resource allocation or the pricing of externalities requires a common metric for costs and benefits. 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, Schelling 1968, Mishan 1971, Jones-Lee 1974), 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, Navrud et al. 2011). Whereas there is consensus about the WTP approach there has been some controversy regarding the empirical elicitation of individual WTP.

The motivation for eliciting individual preferences for risk reductions using the WTP approach to obtain monetary values is that no easily available market prices exist for safety. Instead researchers rely on what is usually referred to as non-market evaluation techniques. These techniques can broadly be classified as either revealed- (RP) or stated-preference (SP) methods. The former refers to methods that use individuals' actual decisions in markets that are related to the good of interest. For instance, property markets have been used to elicit individuals' preferences to reduce the level of air and noise pollution where the relationship between the property prices and pollution levels has been examined (Smith and Huang 1995, Nelson 2008). The RP approach has extensively, especially in the US, been based on labor market data where the compensation workers demand for risker jobs is assumed to reflect their preferences (Viscusi and Aldy 2003). 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. Weaknesses with the RP approach are, though, that markets do not always exist for the good of interest, that the analysts may not be well informed about the decision alternatives individuals face, and that individuals may not be well informed about actual health risks associated with different decisions.

More controversy has surrounded the second type of method, i.e. the SP approach. As the name suggests the SP approach is based on respondents' stated choices in hypothetical market scenarios. There exist a wide range of different SP methods, but the ones that dominate to elicit individual WTP are the contingent valuation method (CVM) and discrete choice experiments (CE) (Bateman, Carson et al. 2002). The general controversy surrounding these methods among economists (and others) is based on the fact that decisions are hypothetical, which means that respondents do not have incentives to be well informed when making their decision and that their stated decision may not reveal how they would act if the decision would have been real. However, despite the criticism of eliciting preferences based on hypothetical scenarios there has been a large increase in the use of SP studies over the past few decades (Carson and Hanemann 2005). The reasons are related to the shortcomings of the RP approach; non-existing markets, market failures, and/or that the analysts may not be well informed about the decision alternatives individuals face. Regarding health risks SP methods have been used to evaluate a wide range of risks, e.g. contaminated water (Adamowicz, Dupont et al. 2011), road safety (Andersson and Svensson 2008), and cancer (Hammitt and Haninger 2010). The SP approach offers flexibility in creating specific markets of interests and allows the analysts to control the decision alternatives.

The aim of this study is to elicit individual preferences to reduce the risk 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, Herman 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. Since food and water safety are attributes with many features that make them candidates for market failures, such as asymmetric information about safety levels, consumers cannot determine the risk before consumption (i.e. it is an experience or credence attribute (Antle 1995)), and consumers' bounded rationality regarding the ability to process risk information, we prefer to use the SP approach. However, it has been shown that individuals in SP studies have difficulties understanding small probabilities and the usefulness of eliciting WTP for risk reductions in SP studies is an area where there has been a lot of debate. Whereas the opponents have based their criticism on empirical results which suggest that individuals are not capable of understanding the risk-dollar tradeoffs presented to them in SP surveys (see e.g. Hausman 2012), the advocates have argued that these results are usually based on bad survey design and that more recent surveys, where the methodology and knowledge among analysts have improved, provide results where the validity of the results often cannot be rejected (Hammitt and Haninger 2010, Carson 2012).

We aim in this study to, in addition to eliciting respondents' preferences, address some methodological issues and the main objectives of this study are to: (i) elicit preference for food and water safety that can be used for policy purposes, (ii) examine whether respondents are able to treat two health variables in CE as separate variables, and (iii) examine the scope sensitivity of respondents' WTP for health improvements. For these purposes we will use data from a Swedish CE study. In CE respondents are asked to choose between different bundles of goods, i.e. a good consisting of several attributes. Compared to the CVM where respondents state their WTP for one good, either a one or a multi-attribute good, CE have some advantages such as being able to extract more information from respondents' answers, e.g. their WTP for the different attributes instead of only one, and to be less prone to strategic bias in the respondents' answers. Regarding our first objective above we take advantage of the former, i.e. the ability of CE to elicit WTP for several attributes, and elicit respondents' WTP to reduce both mortality and morbidity risk, and we examine whether their WTP differs whether the improvement relates to food or water safety.

The methodological objectives of our paper are related to the respondents' ability to reveal their preferences for small changes in risk in a SP study. As to including two health variables in CE studies, we include both morbidity and mortality risk and we examine whether respondents are able to treat them as separate variables or focus on one of them. The final methodological objective, i.e. the test of scope sensitivity, has been extensively examined in the literature with different results (but mostly in CVM studies rather than CE studies). The overall conclusion is that WTP is sensitive to the size of the risk reduction, but not in line with what theory predicts (Hammitt and Graham 1999). 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 baseline risk 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 marginally larger risk reductions) and between sub-samples (whether respondents in a sub-sample with non-marginal higher risk reductions report a higher WTP). The former is the standard test for scope sensitivity in the literature, whereas the latter is the innovation of our study. The higher risk levels are based on the actual levels for road safety for which WTP has been elicited in several studies (Andersson and Treich 2011). By using the baseline level of road safety for which there is much evidence of the range of WTP, the conclusions that we draw based on our results are not only relevant for the elicitation of WTP for food and water safety, but for the elicitation of WTP using SP methods in general.

The paper is structured as follows. In section 2 we relate our paper to a selection of the relevant literature on the valuation of small health risks. Section 3 describes our data collection and shows some descriptive statistics. Section 4 shows our econometric approach outlining the conditional logit and latent class models, whereas results are shown in section 5. Section 6 concludes the paper with a discussion of the results and their place in the literature.

2. Background

Today the WTP approach is widely accepted as the appropriate approach to monetize health risks. Other approaches that have been, and still are, used are the human capital (HC) approach and implicit valuation. In the HC approach the "value of life" is estimated by the individuals' expected lifetime earnings, hence the value reflects the individuals' assumed marked productivity (Mishan 1982). This approach has lost its significance in welfare analysis, however, since the value is not based on individual preferences for safety. Implicit valuation is estimated using information from safety policies. By examining the relationship between the cost and the health effect of the policy, e.g. the number of lives saved, society's assumed implicit value can be estimated (Ashenfelter and Greenstone 2004).[1]

The empirical evidence from the literature on implicit valuation reveals a wide range of estimates, from negative values, i.e. the program saves more resources than it consumes, to values of several billions US$ per avoided death (Tengs, Adams et al. 1996, Viscusi 1998, Sunstein 2002). In general, the highest estimates per avoided death are found for environmental risks, for which the risk levels often are quite small. The large variation may reflect that other objectives than society's preferences are important to policy makers (Carlsson, Kataria et al. 2011) when allocating their resources and can be criticized for not being cost effective. Regarding the latter, this criticism points out that if resources were allocated

[1] Here we use implicit valuation to define valuation based on policy decisions. Implicit valuation is, however, sometimes also used to define the RP approach, since the estimates from those studies are derived from individuals' observed behavior. In this study we always refer to the latter as RP estimates or valuation.

differently, from high to low cost policies, more deaths could be avoided. Ignoring the extreme values, the variation in estimates between types of risks and sectors of society can, however, also reflect that preferences differ between contexts. For instance, a recent meta-analysis of SP studies estimating VSL found that the overall mean VSL from environmental studies was higher compared to the mean VSL from health and traffic related studies (Lindhjelm, Navrud et al. 2011). However, overall the empirical evidence suggests that the risk-dollar tradeoffs from implicit valuation do not necessary reflect individual preferences as estimated in RP or SP studies (Blomquist 2004).

Most of the empirical research on monetizing individuals' preferences for health risks using the WTP approach has been on mortality risks, i.e. these studies have estimated the VSL. Compared with morbidity risk, where the number of endpoints is very large and diverse, there is small variation in the number of endpoints for mortality risk, which may explain the focus on mortality risk in empirical research. Much of this empirical research on estimating VSL has been in the areas of workplace or traffic safety (Viscusi and Aldy 2003, Andersson and Treich 2011) but also in other areas such as general health and environmental risks (see Lindhjelm, Navrud et al. (2011) for a review). Whereas the RP approach, and particularly applied to workplace safety, has dominated in the US, the SP approach has been more applied in Europe and developing countries (Lindhjelm, Navrud et al. 2011). The literature has been dominated by the hedonic pricing approach (Rosen 1974) and the CVM (Mitchell and Carson 1989) using the RP and SP approach, respectively. Recently the CE approach has gained ground, however. In this brief review our main interest is the use of the CE technique to estimate WTP for mortality and morbidity risk.

The CE technique has a relatively long history within marketing and transport economics where it has been used to model demand for new products and modes of transport with different characteristics (Louviere and Hensher 1982, Louviere and Woodworth 1983). Even though it has been used in both health and environmental economics for about two decades (Adamowicz, Louviere et al. 1994, De Bekker-Grob, Ryan et al. 2012) it is only in recent years, as described, that it has become a popular choice to evaluate health risks in both health and environmental economics. Examples of contexts where CE have been used to evaluate health risks are transport (Hensher, Rose et al. 2009), health (e.g. stroke, heart disease, diabetes) (Cameron, DeShazo et al. 2010), avalanches (Rheinberger 2011), contaminated sites (Alberini, Tonin et al. 2007), and contaminated drinking water (Adamowicz, Dupont et al. 2011). Other studies have used the multi-attribute design of CE to examine the effect of context on respondents' WTP (Tsuge, Kishimoto et al. 2005, Alberini and Ščasný 2011).

The results from the CE 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-analyzes of the WTP literature on mortality health risk evaluation have shown that most VSL estimates fall within the range US$ 1 to 10 million (Viscusi and Aldy 2003, Dekker, Brouwer et al. 2011, Lindhjelm, Navrud et al. 2011).[2] These reviews and meta-analyzes are dominated by the HP and CVM methods, but most of the evidence from the CE studies shows similar results as the other techniques.[3] A recent study that found estimates that were outside this range was Adamowicz, Dupont et al. (2011) who found VSL to be in the range C$ 16 to C$ 20 and C$ 14 to C$ 17 million for microbial and cancer, respectively.[4] A difference between their studies and many others in the literature is that they estimated WTP for considerably smaller risks than in the other studies.

In our study we elicit individual preferences for a small risk related to food and water safety. Based on the empirical evidence we can, therefore, expect a relatively high WTP compared to many other estimates, in line with the results in Adamowicz, Dupont et al. (2011). A weakness with any SP study is, as mentioned above, the hypothetical nature of the scenario. Moreover, despite the well-defined ranges of WTP estimates for reducing health risks found in the literature, published estimates have been criticized for publication bias, i.e. unexpected values or values not in line with previous findings are less likely to be published (Ashenfelter and Greenstone 2004, Doucouliagos, Stanley et al. 2012). Further, preference elicitation related to changes in health risks seems to be cognitively demanding for respondents (Carson, Flores et al. 2001). Therefore, since individuals often make decisions based on heuristics (Kahneman,

[2] 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).
[3] 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, Florax et al. (2003) finding evidence that SP studies produce higher estimates than RP studies and Kochi et al. (2006) the opposite.
[4] US$ 1= C$ 1 (stats.oecd.org, 2013-04-07)

Slovic et al. 1982, Kahneman 2003), preference estimates for health risk reductions have been criticized for not reflecting preferences but attitudes (Kahneman, Ritov et al. 1999). In order to examine the robustness of our estimates, and to address the methodological issues related to the hypothetical nature of SP studies, publication bias, and risk comprehension, we take a novel approach and design two alternative scenarios regarding the mortality risk; one using the actual baseline risk to define the change in risk and another using the baseline risk for transport safety. The motivation for using transport safety is because 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). Hultkrantz and Svensson (2012) reported a VSL range equal to USD 0.7 to 8.3 million with a mean and median equal to USD 2.9 and 2 million. We will use the empirical evidence from Sweden on individuals' WTP to reduce road mortality risk to test the robustness of our results in this study.

3. The Survey and Data Collection

In order to address the research questions as set out in the Introduction we administered a CE survey. In the experiment respondents were asked to choose between different public policies that were described to reduce campylobacter-related mortality and morbidity risks. 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.

Preferences and WTP estimates for food and water safety are implicitly derived from the respondents' choices in the CE, which answers our first research question. To address the second research question we test whether respondents can deal with two health variables in CE by examining WTP for the mortality and morbidity attribute, respectively. In addressing the third research question of 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 (smaller risk reduction) and sub-sample B (larger risk reduction), respectively.

3.1 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 CE 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 CE 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.

3.2 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 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 questions, 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 should be irrelevant to respondents' if they only care about the size of the risk reduction, but as showed in some previous research the controllability of a risk may be important in order to understand how individuals perceive and value risks (Slovic 2000). Here we hypothesize that controllability of the risk is lower for water-borne campylobacter and that this may positively affect the valuation of water-borne risk reducing policies.

The mortality risk reductions differed between the two sub-samples since, as mentioned, we wanted to include a substantial test for scope insensitivity. In sub-sample A the mortality risk reductions varied between 1, 2 and 4 fewer deaths per year. This corresponds to reasonable risk reductions 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 sub-sample B the levels were multiplied by a factor of 100 to make the risk reductions in line with road-fatality risk, the risk that has been used in the majority of studies estimating VSL in a Swedish context. 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.[5] Not only does this design permit us to test for scope insensitivity, 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 (Lindhjelm, Navrud et al. 2011, Hultkrantz and Svensson 2012). 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).

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.

[5] 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.

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, Navrud 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) 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 was faced with eight choice sets.

3.3 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".[6]

An example of a choice set, as faced by respondents in sub-sample A, is shown in Figure 1 below. 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.

[Figure 1 about here]

[6] 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 (if becoming sick). We also included a "cheap talk script" in order to mitigate some of the potential hypothetical bias that may arise in SP studies.

After the respondent's first choice he/she was provided some feedback on the computer screen on the meaning of his/her choice regarding changes in risk, the cost associated, etc. The respondent was then asked if he/she was happy with his/her choice and wanted 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 were not given the possibility to change their decisions.

It should be noted that the policies to reduce health risks, both in the water- and food-context, were described as public policies (rather than private individual risk reductions) reducing the risks for the society as a group. The drawback of eliciting "public values" is that they may contain altruistic components that may lead to double-counting of benefits (depending on the type of altruistic preferences (Jones-Lee 1992)). Despite this risk we choose a public scenario, as in e.g. Adamowicz, Dupont et al. (2011) who use a similar approach, since our aim is to obtain "social values" for policy making and also given the risk context in this study we argue that it is substantially more realistic to frame health improvements as a public policy. With this in mind our estimated values may be seen as a "theoretical" upper bound of the true social value of the health improvements.

3.4 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) in 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 2010, SCB 2011). 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 and risk perception of food poisoning in general as well as food poisoning specifically due to campylobacter. 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. 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 (Hakes and Viscusi 2004, Andersson and Lundborg 2007, Andersson 2011), 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, Jendteg et al. 1991, Andersson and Lundborg 2007, Koltowska-Häggström, Jonsson et al. 2007, Andersson, Hammitt et al. 2013)).

[Table 3 about here]

4. Empirical model

4.1 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

$$U_{njt} = \beta_0 sq_{njt} + \beta_1 die_{njt} + \beta_2 sick_{njt} + \beta_3 water_{njt} + \beta_4 cost_{njt} + \beta_5 delay_{njt} + \varepsilon_{njt} \tag{1}$$

where $\beta_0,...,\beta_5$ are coefficients to be estimated, sq_{njt} 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. The remaining attributes in the utility function are described in Table 1 above.

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 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.

Following Viscusi, Huber et al. (2008) we also estimate models in which we interact the *delay* attribute with the *die* and *sick* attributes. This is done to test the robustness of the results and to account for the fact that a delay in the implementation of a policy may cause respondents to lower their valuations of the risk reductions.

4.2 Latent class models

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

$$U_{njt} = \beta_{c0} sq_{njt} + \beta_{c1} die_{njt} + \beta_{c2} sick_{njt} + \beta_{c3} water_{njt} + \beta_{c4} cost_{njt} + \beta_{c5} delay_{njt} + \varepsilon_{njt} \quad (3)$$

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.

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

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

where $V_{njt|c} = \beta_{c0} sq_{njt} + \beta_{c1} die_{njt} + \beta_{c2} sick_{njt} + \beta_{c3} water_{njt} + \beta_{c4} cost_{njt} + \beta_{c5} delay_{njt}$ 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

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

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

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

It should be noted that the number of classes, *C*, must be specified prior to estimating the model. In practice C is unknown, and so a common strategy is to estimate the model with different numbers of classes and choose the preferred specification based on goodness-of-fit measures such as the Akaike and Schwarz criteria. We return to this issue in the Results section.

5. Results

5.1 Conditional logit results

Table 4 presents the result of the baseline model estimated on sub-sample A and B. It can be seen that, everything else equal, respondents prefer policies with lower costs and which lead to greater reductions in the probability of death and illness. The average respondent is also found to have a negative preference for the status-quo alternative.[7] There are no qualitative differences between the sub-samples in terms of the sign and significance of the coefficients.

[Table 4 about here]

[7] 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.

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 3 177 million (95% CI: 2581-3772) (USD 483 million). In comparison, according to the model estimated on sub-sample B the VSL is SEK 50 million (95% CI 31-69) (USD 7.4 million).

As can be seen in Table 4 results for the VSI are very similar in the two sub-samples at SEK 0.34 and 0.33 million (approx. USD 0.05 million), respectively. Hence, when changing the mortality risk reduction between the two-sub samples, we get large effects on estimated VSL whereas we get no statistically significant effect on other attributes such as the VSI. This conclusion is supported by Figure 2, which plots the coefficients in the model estimated on subsample B against the corresponding coefficients in the model estimated on subsample A.[8] The slope of the line in the figure is equal to the relative scale of the two models (Hensher, Louviere et al. 1998). It can be seen that with the exception of the coefficient for mortality risk and the constant for the status-quo alternative, the coefficients in the models are very similar once scale differences are accounted for.

[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. Following Viscusi et al. (2008) we also estimated models in which we interacted the *delay* attribute with the *die* and *sick* attributes. As explained in section 4 this is done to test whether a delay in the implementation of a policy causes respondents to lower their valuations of the risk reductions. We found that the null hypothesis of the interactions being jointly equal to zero could not be rejected at conventional significance levels.[9]

[8] The *sick* and *cost* coefficients have been multiplied by 10,000 and 1,000, respectively, to have a comparable magnitude to the other coefficients.
[9] P-values: 0.24 (sub-sample A) and 0.59 (sub-sample B).

We also explored including interaction effects to allow for observed preference heterogeneity. To be specific, we interacted the attributes with dummies for age group (<35, 35-55, 55+) and income (less/more than SEK 30000 per month). The reference category was a respondent aged less than 35 with an income of under SEK 30000 per month. In general the interactions were found to be individually insignificant, with no clear patterns in terms of preference heterogeneity. Furthermore, the null hypothesis that the interactions are jointly equal to zero could not be rejected at conventional significance levels.[10]

5.2 Latent class model results

Tables 5 and 6 presents the result of a latent class model with 4 classes estimated on the sub-sample A and B. The 4-class models were chosen since they had lower AIC and BIC statistics (better goodness-of-fit) than models with 2 and 3 classes. Furthermore, 4 was the largest number of classes that could be identified with the sub-sample B data.[11] The models were estimated using the EM algorithm with additional Newton-Raphson steps.[12]

The latent class results suggest that there are 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 class of respondents who have very low sensitivity to cost - their cost coefficient is insignificantly different from zero - and a high sensitivity to reductions in the risk of death and illness. 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 group of respondents who have very 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 regardless of the benefits of the proposed policy. Finally,

---

[10] P-values: 0.21 (sub-sample A) and 0.13 (sub-sample B).The results from the models with interactions are available from the authors upon request.

[11] A 5-class model estimated on the sub-sample B data had one class in which all the coefficients had extremely high standard errors, which is a sign of identification problems. Models with more than 4 classes estimated on sub-sample A had better goodness-of-fit than the 4-class model, but did not lead to any additional qualitative insights into the respondents' behaviour (see the discussion below). We therefore decided on reporting models with 4 classes for both datasets to make the results more directly comparable.

[12] To be specific the parameters were estimated by maximising the log-likelihood using the EM algorithm until convergence (convergence was declared when the proportional change in the log-likelihood over the last five iterations was smaller than 0.00001). Additional Newton-Raphson steps were then performed, again until convergence. The estimations were carried out in Stata using the *lclogit* command (Pacifico and Yoo 2013) and code written by the authors. We estimated the models using several different sets of starting values to reduce the chance of the algorithm getting trapped in a local maximum.

there are groups of respondents who trade off the advantages of the policies against their costs, which is more in line with the theoretical expectations regarding consumer behavior.

The finding that the attribute coefficients are insignificantly different from zero in some classes may imply that the respondents ignored some of the information in the experiment, which is in line with the growing literature on attribute non-attendance (Hensher 2010). In the context of the present study it is of particular concern that we find evidence of non-attendance to cost, as this complicates the calculation of VSL. To be specific, individuals who do not take costs into account implicitly have an infinitely high VSL, which is clearly unrealistic and likely to be an artifact of the hypothetical setting rather than a reflection of true preferences. One possibility would be to calculate the VSL based on the preferences of the respondents who did take the cost of the policies into account, but that inevitably raises the question of how representative the estimates are of the population VSL.

Doing the latter we calculate VSL in sub-sample A and B using only class 3 and 4 respondents.[13] This produces a VSL of SEK 1 555 million (95% CI: 897-2 213) (USD 237 million) in sub-sample A and a VSL of SEK 10 million (95% CI: 1-19) (USD 1.5 million) in sub-sample B. In comparison to the conditional logit estimates both VSL estimates are lower, and now with an even higher relative difference between sub-sample A and B. The VSI is estimated to 0.27 million SEK (95% CI: 0.18-0.37) (USD 0.04 million) in sub-sample A, which is relatively close to the conditional logit results. In sub-sample B it is not statistically significant and we therefore refrain from providing an estimate.

[Table 5 about here]

[Table 6 about here]

[13] When the cost coefficient and the coefficients for risk reductions are significant in both classes we report the average VSL/VSI over classes, using the class probabilities as weights. Otherwise we report the individual class estimates for the class in which the coefficients are significant. The individual class estimates are also reported in Tables 5 and 6.

6. Discussion

This study employed a choice experiment to elicit preferences for food and water safety related to campylobacteriosis. One objective of the study was to elicit monetary preference values to be used for policy purposes. However, of major importance were the methodological objectives to examine whether respondents' decisions in CE are affected by only one or both of the health variables, and the scope sensitivity of the respondents' WTP for risk reductions. To answer these research questions we constructed a CE study with attributes including different levels of both mortality and morbidity risk reductions, where respondents also were randomized to one of two sub-samples that differed (by a factor of 100) in the attribute levels of the baseline mortality risk.

The results from our baseline model 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. However, when extending our analysis with latent class models we find evidence of preference heterogeneity and non-attendance, i.e. respondents ignore some of the information in the experiment, which is in line with a growing literature (Hensher 2010). For instance, we find a substantial group of respondents who have very low sensitivity to cost and a high sensitivity to risk reductions, as well as a group who have a very low sensitivity to risk reductions. In the context of CE studies it is of particular concern to find evidence of non-attendance to cost, as this complicates the calculation of monetary values of the attributes. To be specific, in our study individuals who do not take the cost of the program into account implicitly have infinitely high VSL and VSI. In the end we therefore compare our results for VSL and VSI from our baseline model with the latent classes where respondents take into account the cost attribute when answering the choice sets.

Hence, as suggested in the previous paragraph, respondents in our baseline model and a subgroup of the respondents in the latent class model do take into account both the health variables when making their choices. That is, there is evidence of scope-sensitivity (i.e. a significant risk reduction coefficient with the expected sign) in both models. However, even though we found evidence of weak scope-sensitivity, the sensitivity to scope was not adequate across sub-samples which lead to substantially different VSL estimates in our two subsamples. In our baseline logit model VSL was SEK 3 177 million (USD 483 million) and SEK 50 million (USD 8 million) in sub-samples A and B, whereas the results in the latent-class model were SEK 1 555 million (USD 237 million) and SEK 10 million (USD 1.5 million) in sub-sample A and B (Class 4). Hence, with a 100 times smaller risk reduction in sub-sample A the VSL is 60 to 150 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, Navrud et al. 2011). Our results from sub-sample A are at the very high end of the range. Our results from sub-sample B fall very well within the range of previous published estimates, and are close 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 sub-sample analysis suggests no scope sensitivity which questions the suitability of our VSL for policy purposes; the validity of our estimates as reflecting respondents' "true preferences" can be questioned. 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 no 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 morbidity risk. Our robust VSI estimates only strengthen our conclusions regarding our VSL estimates. Moreover, the between sub-sample comparison highlights that even if a study finds weak scope sensitivity based on given choice sets, this does not necessarily suggests that the estimated WTP is a valid measure of individual preferences. As discussed by Goldberg and Rosen (2007) the systematic and repeated questions respondents answer in the CE 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.

Apart from adding and addressing a number of concerns with previous established "ranges" of VSL estimates in the literature and as used in economic policy, 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 and Knetsch 1992, Kahneman, Ritov et al. 1999, Carson 2012, Hausman 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, Solow et al. 1993, p.4607). Some authors have argued that scope insensitivity is avoidable in well conducted CVM (SP) studies and has highlighted that insensitivity to scope has been rejected in many studies (Carson 1997, Carson 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 reflect "moral satisfaction" or expression of attitudes rather than economic WTP for the good/program or attribute (Kahneman and Knetsch 1992, Kahneman, Ritov et al. 1999). 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.[14]

If scope insensitivity is particularly a substantial concern in studies estimating VSL, as has been discussed by others and highlighted in our study, what lessons can be drawn for economic policy and BCA involving effects on mortality? One potential solution for policy evaluations where an estimate of VSL is necessary would be to turn to RP estimates using e.g. wage-risk studies. However, it should be noted

[14] Note that in our study we used frequencies to avoid the difficult cognitive task of evaluating probabilities for respondents.

that RP studies may be plagued by the same type of bias shown in this paper. RP studies assume that individuals have accurate and complete information over risk differences across jobs or consumer products, and that analysts have full information about the consumption alternatives individuals face. There are several reasons to believe that there are systematic misperceptions between objective and subjective risks across jobs and consumer products and that this leads to inconsistent estimates of VSL in RP studies as well (e.g. Hakes and Viscusi 2004). Also, as can be seen in Viscusi and Aldy (2003) and discussed by Viscusi (2012) wage-risk studies may also be sensitive to data issues and estimates from a single country (as in the UK) has been shown to be spread over a large range of values. Hence, in many circumstances RP studies may also provide large variations in estimates or, as is relatively common, are not possible to conduct at all due to lack of data. Moreover, many RP studies have elicited preferences for the same risk contexts, especially studying wage-risk differentials (Viscusi and Aldy 2003), which together with the same risk of publication bias mentioned above may have resulted in the well-defined ranges found for RP data as well.

Despite this uncertainty as to the “true VSL”, proponents of SP methods often argue that a fairly accurate number is better than no number. Critics of this approach, on the other hand, argue that no number may be better than an incorrect estimate from an SP study (see e.g. Hausman 2012). We consider both standpoints equally problematic. On the one hand, if we know that VSL is strictly positive it is hard to justify setting the non-market benefit equal to zero, even if we are concerned about the precision of the estimates. On the other hand, if we are uncertain whether the “true” VSL is USD 8 million or USD 483 million, the uncertainty of the economic evaluation is so large that it is not very informative to a decision maker. We believe that our findings are not only important for future work with CE or any other SP technique, but also different RP approaches such as hedonic pricing. We have shown that standard scope sensitivity tests are not sufficient to test the validity of respondents’ WTP. Moreover, our findings suggest that estimates in line with other studies, which would suggest that 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 examines whether the estimates are in line with other findings in the literature, may be misleading. To summarize, our findings cast doubt on currently suggested policy estimates of VSL based on estimates from the empirical literature, stated as well as revealed preference studies.

To conclude, our results suggest that the standard tests of validity and reliability used in SP and RP studies may not be sufficient to examine whether elicited monetary values reflect individual preferences. This is important due to the large empirical literature on non-market evaluation and the elicited values' central role in policy making, such as BCA.

References

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.

Adamowicz, W., et al. (1994). "Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities." Journal of Environmental Economics and Management **26**(3): 271-292.

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

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. and P. 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, H. and M. Svensson (2008). "Cognitive Ability and Scale Bias in the Contingent Valuation Method." Environmental and Resource Economics **39**: 481-495.

Andersson, H. and N. Treich (2011). The Value of a Statistical Life. The Handbook in Transport Economics. A. de Palma, R. Lindsey, E. Quinet and R. Vickerman. Cheltenham, UK, Edward Elgar Publishing Ltd.**:** 396-424.

Antle, J. M. (1995). Choice and Efficiency in Food Safety Policy, The AEI Press.

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.

Ashenfelter, O. and M. Greenstone (2004). "Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias." American Economic Review Papers and Proceedings **94**(2): 454-460.

Ashenfelter, O. and M. Greenstone (2004). "Using Mandated Speed Limits to Measure the Value of a Statistical Life." Journal of Political Economy **112**(1): 226-267.

Bateman, I. J., et al. (2002). Economic Valuation with Stated Preference Techniques: A Manual Cheltenham, Edward Elgar Publishing.

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

Blomquist, G. C. (2004). "Self-Protection and Averting Behavior, Values of Statistical Lives, and Benefit Cost Analysis of Environmental Policy." Review of Economics of the Household **2**: 89-110.

Brooks, R. G., et al. (1991). "EuroQol: Health-Realted Quality of Life Measurement. Results of the Swedish Questionnaire Exercise." Health Policy **18**: 37-48.

Cameron, T. A., et al. (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., et al. (2011). "Do EPA Administrators Recommend Environmental Policies That Citizens Want?" Land Economics **87**(1): 60-74.

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

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

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 W. M. Hanemann (2005). Contingent Valuation. Handbook of Environmental Economics. K.-G. Mäler and J. R. Vincent. Amsterdam, North-Holland**:** 822-936.

De Bekker-Grob, E. W., et al. (2012). "Discrete Choice Experiments in Health Economics: A Review of the Literature." Health Economics **21**: 145-172.

de Blaeij, A., et al. (2003). "The value of statistical life in road safety: a meta-analysis." Accident Analysis & Prevention **35**(6): 973-986.

Dekker, T., et al. (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.

Doucouliagos, C., et al. (2012). "Are estimates of the value of a statistical life exaggerated?" Journal of Health Economics **31**(1): 197-206.

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

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

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.

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

Hensher, D. A. (2010). Attribute processing, heuristics, and preference construction in choice analysis. Choice Modelling: The State-of-the-art and State-of-practice. S. Hess and A. Daly. Bingley, Emerald Group Publishing Limited**:** 35-70.

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

Hensher, D. A., et al. (1998). "Combining sources of preference data." Journal of Econometrics **89**: 197-221.

Hensher, D. A., et al. (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, L. and M. 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, M. W. (1974). "The Value of Changes in the Probability of Death or Injury " Journal of Political Economy, **82**(4): 835-849.

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

Kahneman, D. (2003). "Maps of Bounded Rationality: Psychology for Behavioral Economics." American Economic Review **93**(5): 1449-1475.

Kahneman, D. and J. L. Knetsch (1992). "Valuing public goods: The purchase of moral satisfaction." Journal of Environmental Economics and Management **22**(1): 57-70.

Kahneman, D., et al. (1999). "Economic preferences or Attitude Expressions?: An Analysis of Dollar Responses to Public Issues." Journal of Risk and Uncertainty **19**(1-3): 203-235.

Kahneman, D., et al. (1982). Judgement under Uncertainty: Heuristics and Biases. Cambridge, Cambridge University Press.

Kochi, I., et al. (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.

Koltowska-Häggström, M., et al. (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, 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.

Louviere, J. and D. A. Hensher (1982). "Design and Analysis of Simulated Choice or Allocation Experiments in Travel Choice Modelling." Transportation Research Record **890**: 11-17.

Louviere, J. and G. Woodworth (1983). "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data." Journal of Marketing Research **20**: 350-367.

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

Mishan, E. J. (1982). Cost-Benefit Analysis. London, UK, George Allen & Unwin, 3rd edition.

Mitchell, R. C. and R. T. Carson (1989). Using Surveys to Value Public Goods: The Contingent Valuation Method. Baltimore, MD., Johns Hopkins University Press.

Nelson, J. P. (2008). Hedonic Methods in Housing Markets: Pricing Environmental Amenities and Segregation. Hedonic Property Values of Studies of Transportation Noise: Aircraft and Road Traffic. Heidelberg, Springer**:** 57-82.

Pacifico, D. and H. Yoo (2013). "lclogit: A Stata command for fitting latent-class conditional logit models via the expectation-maximization algorithm." Stata Journal **13**: 625-639.

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

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

Rosen, S. (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, 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.

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

Smith, V. K. and J.-C. Huang (1995). "Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models." Journal of Political Economy **103**: 209-227.

Taylor, E. V., et al. (2012). "Common source outbreaks of Campylobacter infection in the USA, 1997–2008." Epidemiology and Infection **forthcoming**.

Train, K. (2009). Discrete Choice Methods with Simulation. Cambridge, Cambridge University Press.

Tsuge, T., et al. (2005). "A Choice Experiment Approach to the Valuation of Mortality." Journal of Risk and Uncertainty **31**(1): 73-95.

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. (2012). "WHAT'S TO KNOW? PUZZLES IN THE LITERATURE ON THE VALUE OF STATISTICAL LIFE." Journal of Economic Surveys **26**(5): 763-768.

Viscusi, W. K. and J. 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.

Tables

Table 1 Survey description: attributes and attribute levels

Attribute	Variable name	Attribute levels	
Source of disease	*water*	Food = 0	
		Water = 1	
Mortality reduction	*die*	**Sample A**	**Sample B**
		1	100
		2	200
		4	400
Morbidity reduction	*sick*	8 000	
		16 000	
		32 000	
Delay	*delay*	No delay	
		2 years	
		5 years	
		10 years	
Cost	*cost*	SEK 500	
		SEK 1 000	
		SEK 2 000	

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*
	(1 SEK = USD 6.57, 2012-09-24)	(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
Public Risk perception	Subjective beliefs regarding annual risk of food poisoning (all causes) (objective average risk 10/100)	17.27/100 (17.19/100)	16.73/100 (18.28/100)
Individual Risk perception	Subjective beliefs regarding individual risk of campylobacteriosis per year (average objective risk 7/1000).	16.35/1000 (81.37/1000)	25.65/1000 (118.30/1000)
Health	Health status as measured on a Visual Analog Scale 0-100	80.09 (16.77)	81.94 (15.44)

Note: Standard deviations in parentheses.

Table 4 Benchmark models

	Sub-sample A	Sub-sample B
sq	-0.362***	-0.419**
	(-3.93)	(-2.23)
water	0.230***	0.233***
	(6.01)	(3.08)
delay	-0.105***	-0.0973***
	(-19.15)	(-10.20)
sick	0.0000219***	0.0000167***
	(11.24)	(4.63)
die	0.204***	0.00251***
	(13.18)	(7.85)
cost	-0.000577***	-0.000453***
	(-16.30)	(-6.70)
Estimated VSL[a]	3 177	50
	(2 581 – 3 772)	(31 – 69)
Estimated VSI[a]	0.34	0.33
	(0.27 – 0.42)	(0.16 - 0.51)
Number of respondents	1003	250
Number of responses	8024	2000
Log-likelihood	-8076.01	-1991.46
AIC	16166.02	3996.93
BIC	16200.40	4021.58

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
a:In SEK million. 95 % confidence intervals in parentheses.

Table 5 Latent class model – sub-sample A

	Class 1	**Class 2**	**Class 3**	**Class 4**
sq	2.837**	-3.148***	-4.192***	-0.0552
	(2.15)	(-6.53)	(-13.26)	(-0.21)
water	-0.402	0.296**	0.735***	0.345**
	(-0.86)	(2.34)	(4.32)	(2.27)
delay	-0.0231	-0.227***	-0.143***	-0.123***
	(-0.15)	(-8.60)	(-5.17)	(-4.99)
sick	-0.00000557	0.0000549***	0.00000971	0.0000405***
	(-0.19)	(4.04)	(0.93)	(7.02)
die	0.138	0.259***	0.319***	0.137***
	(0.44)	(3.97)	(3.89)	(2.65)
cost	-0.00164	0.000140	-0.00150***	-0.00133***
	(-1.08)	(1.45)	(-5.98)	(-7.01)
Class probability	0 .187***	0 .335***	0 .307***	0.171***
	(13.94)	(7.08)	(6.62)	(12.43)
Estimated VSL[a]	-	-	1905	927
			(967 - 2843)	(224 - 1630)
Estimated VSI[a]	-	-	Not significant	0.27
				(0.18 – 0.37)
Number of respondents	1003			
Number of responses	8024			
Log-likelihood	-5381.44			
AIC	10824.88			
BIC	10977.11			

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
a:In SEK million. 95 % confidence intervals in parentheses.

Table 6 Latent class model – sub-sample B

	Class 1	Class 2	Class 3	Class 4
sq	-3.753***	3.657	0.985***	-3.803***
	(-5.89)	(1.00)	(2.69)	(-7.15)
water	0.241	1.302*	-0.153	0.954***
	(1.15)	(1.74)	(-0.69)	(2.63)
delay	-0.237***	-0.215**	-0.0838**	-0.134***
	(-6.46)	(-2.09)	(-2.29)	(-3.91)
sick	0.0000393***	0.0000855	0.0000151	-0.00000317
	(3.15)	(1.28)	(1.45)	(-0.22)
die	0.00182***	0.0361***	0.0000699	0.00188*
	(2.61)	(3.32)	(0.10)	(1.93)
cost	0.0000324	-0.000320	-0.000965***	-0.00170***
	(0.21)	(-0.86)	(-4.24)	(-5.48)
Class probability	0.363***	0.120***	0.263***	0.254***
	(6.55)	(4.48)	(9.38)	(4.90)
Estimated VSL[a]	-	-	Not significant	10 million SEK (1 – 19)
Estimated VSI[a]	-	-	Not significant	Not significant
Number of respondents	250			
Number of responses	2000			
Log-likelihood	-1384.22			
AIC	2830.44			
BIC	2939.61			

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

(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?)

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 text in the parenthesis at the top of the figure was only presented to the respondents in the first choice set.

Figure 2 Plot of coefficients in the two models