UNIVERSITY OF OSLO

HEALTH ECONOMICS
RESEARCH PROGRAMME

Valuation of life: a study using discrete choice analysis

Weizhen Zhu
*Department of Economics,
The Ragnar Frisch Centre for Economic Research and HERO*

Working Paper 2004: 3

Valuation of life:
a study using discrete choice analysis

Weizhen Zhu*

May 2003

Health Economics Research programme at the University of Oslo
HERO 2004

* Author's address: Ragnar Frisch Centre for Economic Research, Gaustadalléen 21, N-0349 Oslo, Norway. E-mail: weizhen.zhu@frisch.uio.no

Health Economics Research programme at the University of Oslo
Financial support from The Research Council of Norway is acknowledged.
ISSN 1501-9071, ISBN 82-7756-135-0

Preface

Supervisor: Professor Jon Strand at Department of Economics, University of Oslo.

Abstract[1]

The focus of this paper is to discuss and compare different approaches to calculate the statistical value of life (VSL) based on survey data.

In this paper, we find out that people significantly prefer to reduce the premature death related to the environmental pollution than to reduce the premature death caused by heart disease by using discrete choice technique and estimate a simple logit and ordered logit model. But no significant evidence indicates saving lives from environmental pollution is more preferred than saving lives from traffic accident, or vice versa. VSL is directly calculated from preferences based on our estimates.

We try to link the WTP with the random utility framework in this paper. A new way to make use of the information of WTP is introduced. We show that in theory the common estimates on study of the relationship between WTP and other socio-economic variables by using OLS is biased due to the selection problem. By introducing an "instrument variable" into the regression, it's possible to correct the selection bias.

Acknowledgement[2]

I acknowledge my supervisor Professor Jon Strand, who led me to figure out the ideas of how to write the thesis and gave me fruitful advices on the paper. Many thanks to Professor Olav Bjerkholt for his valuable suggestions on the writing and language, to Professor John K. Dagsvik for his interesting lectures on the discrete choice theory and his inspirations, to Professor Steinar Strøm for his insightful comments and to Tao Zhang and Øystein Børnes Daljord for their useful conversations and kind help with the language. Finally, I thank my husband Zhiyang Jia for his tolerance, encouragement and support through the long process of completing this thesis.

---

[1] This thesis is a part of work under the HERO program at the Frisch Centre for Economic Research.

[2] The responsibility of any errors is absolutely mine.

Table of Contents

1. Introduction

According to Maslow's Hierarchy of Needs Theory (Norwood, 1996), the individual will have higher hierarchy needs when his needs in the lower hierarchy are mostly satisfied. People in a highly developed country such as Norway have already overcome the stage of need for food and clothing. People care more and more about the living qualities such as public health services, traffic safety, and environmental quality, etc. Therefore, the increased risk of severe diseases and premature deaths associated with environmental pollution and traffic accident attracts more and more people's attention.

Each year, in Norway, among about 4.4 million people (Statistics Norway 2002), approximately 19,000 people die due to cardiovascular diseases, approximately 10,000 people die of cancer, and approximately 300 die in car accidents. An unknown number of people die directly or indirectly from different environmental problems (EP), because those problems, both indoors and outdoors, may trigger off or worsen diseases, which cause a premature death.

To achieve the goal of a longer and healthier life requires efforts from both individuals and society. On one hand, individuals can reduce the risk of premature death to some extent by changing their habit, for instance, driving carefully, quitting smoking, exercising more, or moving to a place with less pollution, etc. On the other hand, the government could help to reduce these risks by carrying out a variety of policies such as improving the public health service, conducting road construction, and pollution control etc. It is obvious that the government plays a very important role in this aspect. However, due to the constraints of economic and human resources, the government will not be able to implement every project that will reduce risk of premature death and improve people's life quality. So how to select some projects among a set of possible projects is a practical and also difficult problem. To do this, first of all, we should properly measure, or rank all the possible projects or approaches.

One possible easy way is to set the risk reduction priorities based on the magnitude of the hazard, but this method is not often used. For example, the traffic accident rate in Norway was 9.1/100000, and the Suicide and intentional self-harm rate is 13.1/100000 (Statistics Norway 2002). Obviously, the rate of the latter is much higher than that of the former one, but we know government pays more attention to improve the transportation situation. In this example, the government can do little when people want to give up their own life. Therefore, policymakers consider not only the risk magnitude, but also lots of other factors, for instance, the people’s preference, the degree of difficulty to carry out, and etc.

Generally, most of the justification for the policymaking rests on the cost-benefit analysis. Formal cost-benefit analysis compares the monetary benefits and costs of government actions aimed at improving public welfare. However, because no market price exists for human life, it is generally immeasurable in monetary terms. So it is really difficult, if not impossible, to use this analytical tool for the reduction of premature death.

Partly due to this reason, in the last nearly 40 years, there are lots of discussions about valuation of lives. After all those years debate, researchers began to use willingness to pay (WTP) for a reduction in the probability of death to infer the value society places on saving one anonymous human life. Drèze (1962) noted that the monetary equivalent chosen must reflect the preferences of the individuals affected by the project being evaluated and will thus implicitly involve the trade-off between risk and wealth, while Schelling (1968) first presented the willingness to pay approach in the life saving context. Although WTP is well defined, to emphasize that it means the valuation of a change of risk rate rather than the valuation of the life of a particular individual. In this context, the term Value of Statistical life (VSL) is used. As a matter of fact, it is never possible to value the life of a particular person.

There is a considerable literature using the concept of value of statistical life (VSL). Some examples include here: Dionne and Michaud (2002) analyzed the variability of the value of life estimates; Ghosh et al. (1975) studied value of driving time based on wage rates; Blomquist (1979) did some research for the automobile safety; Portney

(1981) studied environmental health risk; Garbacz(1989) used VSL on housing safety(fire detectors); etc.

The introduction of VSL provides policy makers an easy tool for evaluating different public policy options. In practice, most of the social choices with respect to mortality risks are often made based on the value of a statistical life .If the value of saving one statistical life exceeds the costs incurred, then the project will be worthwhile to undertake.

Intuitively, if a policy reduces the chance of premature death from 7 in one million to 6 in one million, in a population of one million, then that policy is said to save one statistical life. For example, if a project costs NOK 5 million per life saved and we know that VSL is NOK 10 million from some studies, and then we'll see immediately that this policy is worthwhile to implement.

So the question comes down to how to calculate VSL. Rosen (1988, p.287) defines the value of a life as the marginal rate of substitution between wealth and risk. Viscusi (1993) discusses several ways of calculating VSL in different cases. Section 4.4.3 and 6.3 present the detailed calculations of VSL.

There are two main ways to derive VSL values, through revealed preferences or through stated preferences. See example, Morrall (1986), Viscusi (1993), Tengs (1995) and Strand (2002). Revealed preference studies are based on compensating wage data (labor market) or consumer behavior, and stated preference method assesses the value of non-market goods by using individuals' stated behavior in a hypothetical setting, including a number of different approaches such as conjoint analysis, contingent valuation method (CVM) and choice experiments.

Most of the early work on VSL was based on revealed preferences, either based on labor market data or consumer behavior. Among others, Afriat (1972), Hanoch and Rothschild (1972), Diewert and Parkan (1983), and Varian (1984, 1985, 1990) directly applied the revealed preference approach to the production analysis. The problem with the revealed preference approach is that the application of this approach requires some untested assumptions about individuals' risk perceptions. It's often difficult to separate objective risk measures from other subjective attributes of the job or product examined. On the other hand, stated preference studies can normally test whether individuals correctly perceive mortality risks and changes in mortality risks. One of the main advantages of SP approaches is that the analysis need not be constrained by the analysis of market data. Besides the advantages of SP approaches, there are several potential problems with the SP approach to VSL. For instance, one is sensitivity of VSL to the assumed magnitude of risk, or 'scope', where by average stated willingness to pay (WTP) figures per statistical life from stated preference studies have been observed to depend strongly on the magnitude of mortality risk to be valued (Strand 2002). Despite of the drawbacks, the stated preference method seems to be more preferred over revealed preference in the literature (Strand 2002). In the last 20 years, researchers have become more and more in favor of applying stated preference approach. See e.g. Krupnick et al. (2002).

In this paper, we try to recover people's preference over different causes of premature death, and thus estimate VSL using the data from a survey that was conducted in the summer of 1995 in Norway by the Frisch Centre. This survey intended to evaluate some public projects with effect on premature death. The survey provides us data on ranking choices of different projects, and also dichotomous-choice (yes or no questions of whether one is willing to pay the cost in his/her most preferred project) with respect to the first best choice. And after that, we also have open-ended question about how much is the individual's WTP for his/her first preferred choice. This richness of the data provides possibilities of an in-depth analysis of people's behavior and preference.

In this paper, first, we use discrete choice technique and estimate a simple logit and ordered logit model to recover the preference associated with ranking and the preference associated with the risk reduction by using the ranking data (section 4). The results from these two models are quite similar, which may be seen as indicator of quite good data quality. We find out that people significantly prefer to reduce the premature death related to the environmental pollution than to reduce the premature death caused by heart disease. But no significant evidence indicates that saving lives from environmental pollution is more preferred than saving lives from traffic accident, or vice versa.

We also calculate the VSL directly from preferences based on our estimates. But the VSL found here is a bit high compared with other studies. This agrees with the findings in Halvorsen and Sælensminde (1998). They claimed that individuals react differently to a dichotomous-choice CVM question than to a ranking one. In her paper, Halvorsen (2000) used the much more sophisticated technique of nested logit. Here we place the dichotomous-choice answers into a simple ranking framework instead, and use an approach that is less technical and easier to understand to elicit that the ranking and the dichotomous-choice are not consistent. So using results from the logit model to calculate VSL may not be appropriate.

Another widely used approach for calculating VSL is to use WTP regressions. Namely, researchers often do regression of WTP on some variables of interest, such as income, age, education and so on. One problem with this approach is the so called 'selection bias' problem, which arises since only the WTP for first best choice are observable. To solve this problem,, we try to link the WTP with the random utility framework in this paper. We suggest a new way to make use of the information of WTP. By introducing an “instrument variable” $z = \ln(P_n^*(i))$ into the regression, we can succeed to correct the selection bias. Where $P_n^*(i)$ is the predicted probability of the chosen project i. Essentially, this is similar to the well-known ‘Heckman two step method’ (Section 6).

We show that in theory the common estimates on study of the relationship between WTP and other socio-economic variables by using OLS is biased due to the selection problem. And our preliminary study shows that danger of ignoring the selection problem does exist when we compare the empirical results from these two methods.

The rest of this paper is organized as follows: Section 2 presents the methods we’ll use in this paper, namely contingent valuation method and choice experiment. Section 3 reviews the survey and data descriptions. In section 4, we present the theoretical model settings for the simple logit and ordered logit, and analyzed the empirical findings from the ranking data. Section 5 presents a new way to utilize the dichotomous-choice information. In section 6, we develop a two-step model to relate the WTP and random utility framework. Section 7 concludes.

2. Contingent valuation method and choice experiments

In the survey, from which we get our data, when the respondents were asked to make choices, the choice experiment method occurs. And furthermore asking the respondents to state their WTP involves contingent valuation method. So, the dataset we use here is from some a combined choice-experiment and contingent-valuation survey.

In this section, we'll briefly introduce the contingent valuation method and choice experiments method. In our case, these two methods complement each other.

2.1 Contingent valuation method

There is a dichotomous-choice question and an open-ended question in our survey. This survey method is a sort of contingent valuation method (CVM). It is called 'contingent' valuation, because people are asked to state their willingness to pay, contingent on a specific hypothetical scenario and description of the environmental service.

The contingent valuation method (CVM) asks people to state their values directly, rather than inferring values from actual choices. For instance, people might be asked to state their maximum willingness to pay (WTP) for some environmental service or to state their minimum to accept compensation (WTAC). So CV is a 'stated preference' method, rather than 'revealed preference' method. Like the other SP methods, CVM analysis need not be constrained by the analysis of market data. And furthermore CVM is a direct stated preference method. The directness is one of the main strengths of CVM. It results in a single understandable measure, which is expressed in monetary term. This value from CVM is able to capture many of the 'externalities' of environmental and cultural resources, compared with other indirect stated preference method.

Since the WTP question is asked directly, CVM normally uses relatively simple questionnaire formats. This simplicity strength might be a bit obscure, but generally it enables respondents understand the questions more intuitively.

There are some weaknesses of CVM as well. The most known one is that the value from CVM is likely to be strategically biased, since the individuals interviewed have incentives to answer untruthfully. For example, if the individual has to pay the amount equal to he/she statement, then he/she may have incentives to understate. This is a type of free-riding problem. Some other problems such as question framing, and scenario misspecification are also disadvantages of CVM. Halvorsen et al. (1996) discussed detailed strengths and weaknesses of this method.

The contingent valuation method (CVM) is used to estimate economic values for all kinds of environmental services. It can be used to estimate both use and non-use values. In most applications, CVM has been the most commonly used approach, although it's also a very controversial approach. The idea of CVM was first suggested by Ciriacy-Wantrup (1947), and the first study was in 1961 done by Davis (1963)(An economic study of the Maine woods). Mitchell and Carson (1989) give detailed overview of CVM.

Any CVM exercise can be split into five stages: (1) setting op the hypothetical market, (2) obtaining bids, (3) estimating mean WTP and/or WTAC, (4) estimating bid curves, and (5) aggregating the data. See Hanley, et al. (1997).

2.2 Choice experiments

The results of contingent valuation surveys are often highly sensitive to what people believe they are being asked to value, as well as the context that is described in the survey. Thus, it is essential for CVM researchers to clearly define the services and the context, and to demonstrate that respondents are actually stating their values for these services when they answer the valuation questions. So at the same time as CVM was developed, other types of stated preference techniques, such as choice experiments (CE), evolved in both marketing, transport economics and lately environmental economics.

In a choice experiment, individuals are given a hypothetical scenario and asked to choose their preferred alternative among several alternatives in a choice set, and they are usually asked to perform a sequence of such choices. Each alternative is described by a number of attributes or characteristics. A monetary value is included as one of the attributes, along with other attributes of importance, when describing the profile of the alternative presented. Thus, when individuals make their choice, they implicitly make trade-off between the levels of the attributes in the different alternatives presented in a choice set (Alpizar et al.2001). In our survey, the respondents were asked to make choices among four alternatives. Each alternative includes four attributes: the number of lives saved, the time until effect, the cost, and the causes of premature death.

CE is a method evaluating the preferences of individuals for the relevant attributes of some goods. Therefore, when we need to identify and evaluate different attributes of a good, CE is a good method. Furthermore, there are several more advantages in choice experiments, compared with CVM: (i) reduction of some of the potential biases of CVM, (ii) more information is elicited from each respondent compared to CVM and (iii) the possibility of testing for internal consistency (Alpizar et al.2001).

As a new evaluation technique, CE is a multi-attribute, preference-elicitation technique that is widely used in marketing research and transportation (Louviere, et al., 2000). The first study to apply choice experiments to non- market valuation was Adamowicz et al. (1994). Since then, CE began to be used in environmental economics, such as Boxall et al. (1996, hunting), Hanley et al. (1998, environmentally sensitive areas), Garrod and Willis (1998, landfill waste disposal), Rolfe et al. (2000, tropical forest), Carlsson and Martinsson (2001, donations for environmental projects), Blamey et al. (2000 green products), Layton and Brown (2000, applications to environment), Ryan and Hughes (1997, applications to health) and etc.

3. The data

The data is taken from a survey conducted in the summer1995 by the Frisch Centre. In the survey, 1002 individuals were randomly selected from the whole population of Norway. And they were only asked questions related to this survey. Approximately, the response rate of this survey is 68 percent.

This survey intended to evaluate some public projects and to recover the Norwegian people's preferences for the reduction of premature death caused by three different causes. The survey consists of 9 parts and 24 questions in all. In this paper I will mainly use the answer to question 3 in part 2 and question 4 in part 3.

In question 3, each respondent was asked to make two choices (first and second preferred) between four different projects of reducing the number of people suffering a premature death. The attributes which distinguish the four projects are: i) The number of lives saved by the project; ii) The time lag from when the project is initiated till it starts to save lives; iii) The causes of death; and iv) The annual cost for the respondent's family: the amount which the household would have to pay in terms of higher direct and indirect taxes in order to carry out the project.

One of the survey questions (question 3) is:
If the government chooses project A (B, C, D), they will save _ (note) lives every year after a time lag of _(note) years, who would otherwise have died of _(note). The increase in the direct and indirect taxes necessary to finance this project will cost your family _(note) NOK every year.

> *Question 3 (a):*
> *If the government must choose one of these projects, which one of these four projects do you prefer? A/B/C/D*
> *Question 3 (b):*
> *If the government does not choose the project you ranked as first best, which one of the three remaining projects do you prefer? A/B/C/D*

There were four variations in attributes 1), 2), 4) respectively and three variations in attribute 3). So, in all there were 192 possible combinations of the attributes, which can describe a particular project. The survey designers used an iterative optimizing procedure in SAS called OPTEX and applied an A-optimality criterion to choose 56 combinations. And the survey contains fourteen sub-samples by these 56 combinations. See Halvorsen (2000).

The respondents were asked two CVM questions conditional on the most preferred project after they made the ranking choices. At first, the respondents were asked whether they would be willing to pay the cost of the project they ranked as the most preferred. This is a dichotomous-choice question. Then, they were asked an open-ended question to state their maximum willingness to pay (WTP) for their first choices. Halvorsen (2000) utilized this information by using a nested logit model. In this paper we construct a new structure to fully utilize this information. The question sample is as following:

Question 4:
Now, assuming that the government will carry out the project you preferred in the last question. That is, project __(note). The government will finance the project through an increase in both direct and indirect taxes that will cost your family __(Note) NOK in additional yearly expenses.

> *(a): When you consider your household's annual income and fixed expenditures, are you willing to pay this cost so the government may achieve this project? Remember that this will leave you with less money for i.e. food clothing, shoes, travel, car use and savings.*
> *Yes/No/Don't know.*
>
> *(b):When you consider your household's annual income and fixed expenditures, what is the maximum cost you would be willing to pay so the government may achieve this project? Remember that this will leave you less money for i.e. food, clothing, shoes, travel, car use and savings. ___NOK.*

At the end of the survey, there are several questions designed to collect socio-economic background information. And also there are some questions about how they react towards the question 3 and 4.

One thing to note here is that, there are some missing values in the data set for the first or second choices. But altogether there are only 22 observations without the choices for first or second or both, considering the sample size 1002, I don't think it will affect the main result, so I just drop all those 22 observations. That is, there are 980 observations in the data set that I use.

4. Preference related to ranking – analysis using discrete choice model

In question 3, the respondents were asked to evaluate four different government projects characterized by four attributes: the number of lives saved, time lag until effect, the cost, and the death causes. Obviously, this is a contingent ranking problem. We would like to recover people's preferences of different projects and their preferences of reducing premature deaths of three different causes of diseases.

Simply speaking, the problem facing the respondent is just a discrete choice problem over the proposed projects.

To explain the choices made by the respondent we will employ two different, but related, models. In both of the models, choice probabilities are derived from a classical utility maximizing framework. We will only use just first best choice (the most preferred choice) in the first model. The obvious weak point for this model is that it fails to fully exploit the data given the fact that we not only have the observations on the first best choice, but also data on the second best choice. It could be a waste of the valuable information.

We can use available information on the second best choices in two different ways. First, we can use the information to do some 'out of sample' prediction to see how well the logit model above performs. To do so, we will simply use the estimated parameters from the logit model to predict the second choice and compare it with the actual observations. The second, a more direct method to make use of the second choice observations, is the so-called ordered logit model, which specify the joint probability of ranking alternatives (In our case, we only need to specify the joint probability of the first two choices). This method should provide more precise information about preferences than the simple logit model that relies solely on the highest ranked alternatives.

4.1 The simple logit model:

Similar to all the other discrete analysis, our analysis is based on the random utility framework as well.

From one particular respondent's point of view, the utility is deterministic, but in practice, one may observe that observationally identical respondents make different choices. Thus, there must exist some unobservable factors affecting the individuals' behavior to an econometrician. So from an economist's point of view, the utility is random. For reasons why the utility from the analyst's point of view should be best viewed as a random variable, see Ben-Akiva and Lerman (1985).

So suppose the utility for the respondent (decision-maker) *n* to choose alternative *i* is U_{ni}. Then we can write:

$$U_{ni} = u_{ni} + \varepsilon_{ni}, \quad \text{where} \quad i = 1,2,3,4 \tag{1}$$

Where u_{ni} are the systematic or deterministic components of the utility, and it depends on the attributes of alternative i such as the number of the lives saved, the lagged time. Note that the u_{ni} in general differs from respondent to respondent. ε_{ni} are disturbances or random components. These variables account for unobserved attributes of the states that affect preferences, unobserved taste variation across the respondents. We assume the disturbances ε_{ni} are extreme value distributed random variables just for analytic convenience. Furthermore, we assume that they are IID (independently and identically distributed) across the alternatives and respondents. And also we assume that ε_{ni} is Gumbel-distributed with location parameter η and a scale parameter σ>0. Then we can rewrite the random utility function as:

$$U_{ni} = u_{ni} + \eta + \frac{\varepsilon_{ni}'}{\sigma} \tag{2}$$

Where ε_{ni}' is extreme value distributed with parameter (0,1).

Since each systematic utility has a constant term, we can assume a constant η for all alternatives or $\eta = 0$, which is not in any sense restrictive. Here for convenience, we can just assume $\eta = 0$. Then, equation (2) reduces to:

$$U_{ni} = u_{ni} + \frac{\varepsilon_{ni}{}'}{\sigma} \tag{3}$$

Under the utility maximization assumption, the probability of alternative *i* to be chosen by the decision-maker (respondent) *n* is:

$$\begin{aligned} P_n(i) &= \Pr(U_{ni} \geq \max_{k \neq i} U_{nk}) \\ &= \Pr(u_{ni} + \frac{\varepsilon_{ni}{}'}{\sigma} \geq \max_{k \neq i}(u_{nk} + \frac{\varepsilon_{nk}{}'}{\sigma})) \\ &= \Pr(\sigma u_{ni} + \varepsilon_{ni}{}' \geq \max_{k \neq i}(\sigma u_{nk} + \varepsilon_{nk}{}')) \end{aligned} \tag{4}$$

Let:

$$v_{ni} = \sigma u_{ni} \tag{5}$$

Define:

$$U^*_{n(-i)} = \max_{k \neq i}(v_{nk} + \varepsilon'_{nk}), \tag{6}$$

Then following the property of the extreme value distribution (Ben-Akiva and Lerman (1985)), $U^*_{n(-i)}$ is also extreme value distributed with parameter ($\ln(\sum_{k \neq i} \exp(v_{nk}))$, 1).

Then we can write $U^*_{n(-i)} = v^*_{n(-i)} + \varepsilon^*_{n(-i)}$, where $v^*_{n(-i)} = \ln(\sum_{k \neq i} \exp(v_{nk}))$, and $\varepsilon^*_{n(-i)}$ is extreme value distributed with parameter (0,1).

So we have:

$$\begin{aligned}P_n(i) &= \Pr(v_{ni} + \varepsilon_{ni}{}' > \max_{k \neq i}(v_{nk} + \varepsilon_{nk}{}')) \\ &= \Pr(v_{ni} + \varepsilon_{ni}{}' > v^*_{n(-i)} + \varepsilon^*_{n(-i)})) \\ &= \Pr(\varepsilon_{ni}{}' - \varepsilon^*_{n(-i)} > v^*_{n(-i)} - v_{ni}) \\ &= \frac{1}{1 + \exp(v^*_{n(-i)} - v_{ni})} \\ &= \frac{\exp(v_{ni})}{\sum_k \exp(v_{nk})}\end{aligned} \tag{7}$$

Using the definition of v_{ni} (see (5)), we can write (7) as:

$$P_n(i) == \frac{\exp(v_{ni})}{\sum_k \exp(v_{nk})} = \frac{\exp(\sigma u_{ni})}{\sum_k \exp(\sigma u_{nk})} \tag{8}$$

It is very clear by now that we are only able to estimate v but not u, and there is no way to identify the scale parameter σ. So those parameters, which enter linearly into the utility function, cannot be identified.

Let Y_{ni} denote respondent *n* to choose project *i*, and

$$Y_{ni} = \begin{cases} 1, & \text{if project } i \text{ is chosen} \\ 0, & \text{if otherwise} \end{cases} \tag{9}$$

Denote the probability for respondent n to choose project i is

$$Q_{ni} = \Pr(Y_{ni} = 1) = \Pr\left(U_{ni} \geq \max_k(U_{nk})\right) = P_n(i) \tag{10}$$

The likelihood function is:

$$L = \prod_{n=1}^{N} \prod_i Q_{ni}^{Y_{ni}} \tag{11}$$

So it's obvious the log likelihood function is:

$$
\begin{aligned}
Logl &= \sum_{n=1}^{N}\sum_{i}\left[Y_{ni}\ln(Q_{ni})\right] \\
&= \sum_{n=1}^{N}\sum_{i}\left(Y_{ni}\ln\left(\frac{\exp(v_{ni})}{\sum_{k}\exp(v_{nk})}\right)\right)
\end{aligned} \tag{12}
$$

Where i refers to project i (i=A; B; C; D).

This simple logit only uses the first best choice of the respondents. The advantage of this is that it's very simple, but the weak point is that it doesn't fully exploit the dataset. That is it doesn't efficiently use the useful information. So we introduce ordered logit, which uses both first and second choices, in next section.

4.2 The ordered logit model.

Given the utility structure **(1)**, in the previous section we have already shown that the probability for choosing i as first choice is:

$$
\begin{aligned}
P_n(\text{first choice}=i) &= \Pr\left(U_{ni} \geq \max(U_{nk})\right) \\
&= \frac{\exp(v_{ni})}{\sum_{k}\exp(v_{nk})} \\
&= \frac{\exp(\sigma u_{ni})}{\sum_{k}\exp(\sigma u_{nk})}
\end{aligned} \tag{13}
$$

For the respondent *n*, the probability for choosing *i* as first choice and *j* as second choice is:

$$
\begin{aligned}
&P_n(\text{first choice}=i;\ \text{second choice}=j)\\
&=\Pr\left(U_{ni}\geq U_{nj}\geq \max_{k\neq i}\left(U_{nk}\right)\right)\\
&=\Pr\left(U_{nj}\geq \max_{k\neq i}\left(U_{nk}\right)\right)-\Pr\left(U_{nj}\geq \max_{k}\left(U_{nk}\right)\right)\\
&=\frac{\exp(v_{nj})}{\sum_{k\neq i}\exp(v_{nk})}-\frac{\exp(v_{nj})}{\sum_{k}\exp(v_{nk})}\\
&=\frac{\exp(\sigma u_{nj})}{\sum_{k\neq i}\exp(\sigma u_{nk})}-\frac{\exp(\sigma u_{nj})}{\sum_{k}\exp(\sigma u_{nk})}
\end{aligned}
\tag{14}
$$

Let Y_{nik} denote the respondent *n* to choose project i as first choice and project *k* as second choice.

Suppose:

$$
Y_{nik}=\begin{cases}1, & \text{if choose } i \text{ as first choice, } k \text{ as second choice}\\ 0, & \text{if otherwise}\end{cases}
\tag{15}
$$

Denote $\tilde{Q}_{nik}$ as the probability for respondent *n* to choose project i as first best choice and project *k* as second best choice,

$$
\begin{aligned}
\textit{i.e.}\quad \tilde{Q}_{nik}&=\Pr(U_{ni}\geq U_{nk}\geq \max_{r,q\neq i,k}(U_{nr},U_{nq}))\\
&=\Pr(Y_{nik}=1)
\end{aligned}
\tag{16}
$$

Where i,k refers to the project (A; B; C; or D).

Then the likelihood is:

$$
L=\prod_{n}^{N}\prod_{i\neq k}\tilde{Q}_{nik}{}^{Y_{nik}}
\tag{17}
$$

The log likelihood function will be:

$$
Logl=\sum_{n=1}^{N}\sum_{k\neq i}\left(Y_{nik}\ln(\tilde{Q}_{nik})\right)
\tag{18}
$$

Where i,k refers to the project (A; B; C; or D).

4.3 Specifications of the utility function

Up to this point in our discussion we have not imposed any functional form on u_{ni}, the systematic component of the utility function. It's generally computationally convenient to restrict u_{ni} to the class of linear-in-parameters functions. Even in linear-in-parameters system, u_{ni} can have all kinds of different specifications. Let X_{ni} be the vector of the attributes of project *i*, and x_{nik} refers to the element k in the vector X_{ni}, then, $u_{ni}(X_{ni})$ can e.g. be any one of the following specifications:

$$u_{ni}(X_{ni}) = \sum_k \alpha_k x_{nik} \tag{19}$$

$$u_{ni}(X_{ni}) = \sum_k \alpha_k \ln(x_{nik}) \tag{20}$$

$$u_{ni}(X_{ni}) = \sum_{k \in C_1} \alpha_k x_{nik} + \sum_{j \in C_2} \alpha_j \ln(x_{nij}) \tag{21}$$

where $C_1 \cup C_2$ is the entire set of attributes' index, and $C_1 \cap C_2 = \varnothing$

$$u_{ni}(X_{ni}) = \sum_k \alpha_k \sqrt{x_{nik}} \tag{22}$$

$$u_{ni}(X_{ni}) = \sum_k \alpha_k x_{nik} + \sum_k \beta_k x_{nik}^2 \tag{23}$$

$$u_{ni}(X_{ni}) = \sum_k \alpha_k x_{nik} + \sum_k \beta_k x_{nik}^2 + \sum_k \gamma_k x_{nik}^3 \tag{24}$$

$$\dots$$

Here for simplicity, we use the specification (19). That is, we suppose the utility is linearly related to the attributes.

So we write the utility for respondent *n* choosing project *i* as

$$U_{ni} = \beta + \beta_{\mathrm{TA}} * D_{ni}^{TA} + \beta_{\mathrm{CD}} D_{ni}^{CD} + \beta_t t_{ni} + \beta_l * l_{ni} + \beta_c * (y_n - c_{ni}) \tag{25}$$

Where t_{ni} is the time lag of project *i*, l_{ni} is the number of life saved of project *i*, c_{ni} is the household cost, $y_n - c_{ni}$ is the household disposable income after the respondent n pays the cost.

Note here β is the constant for causes of environmental problem (EP), since it's the same for one individual, we will not be able to identify it. Meanwhile since the utility property won't change if we just subtract a constant β , for simplicity, we can do a normalization, by letting β =0. That is, EP is reference point, and the difference between EP and traffic accident (TA) is β_{TA} , and the difference between EP and cardiovascular disease (CD) is β_{CD} .

Now let's look into specification (25) more deeply. Implicitly, it is assuming that the utility function has different intercepts for the three death-causes, but has the same slope for all three different death causes. The different intercepts means that the utility function's starting point is different. However, the unique slope indicates that the marginal rate of substitution between the number of lives saved and the cost. In practice, this is not likely to be the case. Intuitively, when people consider the choices, they will consider the fact that different death causes affect different groups of people. For instance, Traffic accident normally happens to younger group than the cardiovascular diseases do. Generally, to gain the same utility, it will be different between saving one life from one type of death and from another type of death if it costs the same, by holding time constant. This means that the marginal rates of substitution between the number of lives saved and the cost are different for different death causes. So the utility function should have different slopes for each death cause. To implement this, we can allow the dummy variables to interact with the life variables, that is, the utility function can be specified as follows:

$$\begin{aligned}U_{ni} = \beta + \beta_{\mathrm{TA}} \cdot D_{ni}^{TA} + \beta_{\mathrm{CD}} \cdot D_{ni}^{CD} + \beta_t \cdot t_{ni} + \beta_l \cdot l_{ni} + \beta_{lTA} \cdot D_{ni}^{TA} \cdot l_{ni} \\ + \beta_{lCD} \cdot D_{ni}^{CD} \cdot l_{ni} + \beta_c \cdot (y_n - c_{ni})\end{aligned} \tag{26}$$

Where *i=A, B, C, D.*

In(26), both intercept and slope depend on the dummy variables.

From above discussion, intuitively, specification (26) is more reasonable than specification (25), but is it true? We'll try to find some empirical evidence to support this argument in section 4.4.

4.4 Estimation

For the multinomial logit model, the most widely used method is the maximum likelihood (ML) method. Although there are still other methods that can be applied to a logit model, such as least squares, it has no theoretical advantage over maximum likelihood.

Simply stated, a maximum likelihood estimator is the value of the parameters for which the observed sample is most likely to have occurred. Although maximum likelihood estimators are not in general unbiased, they are consistent and asymptotically normal and efficient. So we can apply asymptotic t test to test whether a particular parameter in the model differs from some known constant, and the likelihood Ratio (LR) test to test some linear constraints of the parameters. Among all the estimation results reported in this paper, we also include two informal goodness-of-fit measures ρ^2 and $\overline{\rho}^2$.

$$\rho^2 = 1 - \frac{\ell(\widehat{\beta})}{\ell(0)}$$

$$\overline{\rho}^2 = 1 - \frac{\ell(\widehat{\beta}) - K}{\ell(0)}$$

where $\ell(\widehat{\beta})$ is the value of the log likelihood at its maximum

$\ell(0)$ is the value of the log likelihood when all the parameters $= 0$;

K is the number of parameters

4.4.1 The simple logit

Recall in section 4.1, we have the log likelihood function: (12). Next we will use two different utility specifications to estimate this simple logit model.

a). Specification (25)

Use the specification(25):

$$U_{ni} = \beta + \beta_{\mathrm{TA}} * D_{ni}^{TA} + \beta_{\mathrm{CD}} D_{ni}^{CD} + \beta_t t_{ni} + \beta_l * l_{ni} + \beta_c * (y_n - c_{ni}),$$

We get the results in Table 1.

Table 1. Estimate for the simple logit model, using the utility specification (25)

Variable	Coef	Estimate	T-value
Dummy causes of TA	β_{TA}	-0.2443	-1.9294
Dummy causes of CD	β_{CD}	-0.3553	-3.4059
Time until effect	β_t	-0.0647	-11.7179
Number of lives saved	β_l	0.0031	9.2928
Cost (in 1000 NOK)	β_c	-0.1816	-6.7273
# of observations		980	
Log-likelihood		-1268.8700	
ρ^2		0.0660	
$\bar{\rho}^2$		0.0623	

Note: TA is traffic accident, CD is cardiovascular disease (heart disease), and EP is environmental pollution. Standard Errors computed from analytic second derivatives.

From Table 1 we notice that the coefficient associated with TA is significant at the 10% level of significance (LOS) in the simple logit model, but it is not significant at the 5% LOS. The coefficients associated with all the other variables have the expected sign and are relatively sharply determined.

The coefficient for the number of lives saved is significantly positive, which means that the utility of choosing a project increases in the number of lives saved of the project, as we expected. And it's not difficult to see, the utility of choosing a project decreases in the time lagged and the cost of the project.

Since the constant for the environmental pollution is the reference point. We know immediately from the results that people have significantly high preference for reducing the premature death related to the environmental pollution, relative to reducing the premature death caused by heart disease, when holding all the other attributes constant. We can also say that people might slightly prefer to save lives from the environmental pollution than to save lives from traffic accident by holding all the other attributes constant, if we use 10% LOS. But if we use 5% LOS, then we can not tell people's preference difference of saving lives from environmental pollution and saving lives from traffic accident.

These results are not surprising though. For most of people, the environmental pollution related premature death is quite 'mysterious'. This 'unknown' property of environmental death is scaring in some sense. So, people's intending to reduce premature death related to environmental pollution is reasonable.

According to the scenario description of the survey, we know that approximately 19000 Norwegians die every year due to heart diseases, and 300 die in traffic accident. From these numbers, we can easily see that a reduction of 100 death is better for traffic accidents than for heart disease because this reduces traffic deaths by relatively much more. And if we investigate it deeply, we find out that most of the people who die of heart disease are old. According to the Statistical Yearbook of Norway 2002, the average age of death caused by heart disease is approximately 70, and the average age of death due to traffic accident is only about 30. So considering the remaining life years saved, it's understandable that people have higher utility to save 1 life from traffic accident than to save 1 life from heart diseases. And traffic accident is also unpredictable, which is increasing the scare of it.

b). Specification (26)

As discussed in section 4.3, we know it is incomplete to assume that the marginal rate of substitution is the same for all the three different death causes. So, here use the specification(26):

$$\begin{aligned} U_{ni} &= \beta + \beta_{\mathrm{TA}} \cdot D_{ni}^{TA} + \beta_{\mathrm{CD}} \cdot D_{ni}^{CD} + \beta_t \cdot t_{ni} + \beta_l \cdot l_{ni} + \beta_{lTA} \cdot D_{ni}^{TA} \cdot l_{ni} \\ &\quad + \beta_{lCD} \cdot D_{ni}^{CD} \cdot l_{ni} + \beta_c \cdot (y_n - c_{ni}) \end{aligned}.$$

Allowing the dummy variables to interact the life variable, we assume that the marginal rate of substitution of lives saved and cost differs between death causes. The results from this specification are in Table 2.

Table 2. Estimates from simple logit model, using the utility specification (26)

Variable	coef	Estimate	T-value
Dummy, causes of TA	β_{TA}	-0.5486	-1.2162
Dummy, causes of CD	β_{CD}	0.0937	0.3137
Time until effect	β_t	-0.0675	-11.9670
Number of lives saved	β_l	0.0091	2.4436
Dummy CD*life	β_{lCD}	-0.0059	-1.5926
Dummy TA*life	β_{lTA}	0.0034	0.6207
Cost (in 1000 NOK)	β_c	-0.1765	-6.4687
Number of observations		980	
Log-likelihood		-1263.1300	
ρ^2		0.0702	
$\bar{\rho}^2$		0.0651	

From results in Table 1, Table 2, both coefficients associated with the interaction on life are insignificant, but are they equal to 0 simultaneously ($\beta_{lCD} = \beta_{lTA} = 0$)? To test this, we can apply likelihood ratio (LR) test to compare these two estimations.

When we got the estimates of both the restricted and unrestricted parameters vectors, normally we can use likelihood ratio test. Let's use our case to illustrate this. The estimation results in Table 1 are the estimates from model with constraints, which is all the coefficients of the interaction on life equal to 0, while the results from are the estimates from unrestricted model. Suppose that the likelihood function values at these estimates of the restricted and unrestricted model are respectively: $\hat{L}_R$ and $\hat{L}_U$.

Define the likelihood ratio as: $\lambda = \dfrac{\hat{L}_R}{\hat{L}_U}$.

The formal test is based on the following result.

Theorem (Greene (2000), pp.152, Theorem 4.20):

Distribution of the likelihood ratio test statistics. *Under regularity, the large sample distribution of* $-2\ln\lambda$ *is chi-squared, with degrees of freedom equal to the number of restrictions imposed.*

According this theorem, we can get the test statistic is:

$$LR = -2\ln\lambda = -2\ln(\frac{\hat{L}_R}{\hat{L}_U}) = -2(\ln(\hat{L}_R) - \ln(\hat{L}_U)).$$

Which is χ^2 distributed by (Ku-Kr) degree of freedom, where Ku and Kr are the numbers of estimated coefficients in the unrestricted and restricted models, respectively. In our case, there are 2 more parameters enter in the latter model. So for the null hypothesis that all the coefficients associated with the interaction on life are 0, the LR statistic will be χ^2 distributed by 2 degree of freedom.

Remember that in the tables we got the log likelihood, so we have

$$LR = -2(\text{-}1268.87 + 1263.13) = 11.48.$$

The corresponding p value<0.01. It means that we can reject the null hypothesis at very low level of significance. That is the interaction part makes difference in the model. This result agrees with our intuitive understanding. So we'll use the specification (26) for the utility in the rest of this paper.

c). Prediction

Use the estimates from Table 2, we can predict the probability for first choice and second choices respectively, we get:

Table 3. Comparison of the observed and predicted first choice proportions.

	Observed First choice's proportions	Predicted first choice's proportions
Death caused by CD	0.612245	0.61224
Death caused by TA	0.22449	0.22449
Death caused by EP	0.163265	0.16327

Table 4. Comparison of the observed and predicted second choice proportions.

	Observed second choice's proportions	Predicted second choice's proportions
Death caused by CD	0.604082	0.61456
Death caused by TA	0.240816	0.21164
Death caused by EP	0.155102	0.1738

From the above two tables, we see that the prediction for the first choice is very good. There are some differences between the predicted and observed proportion for the second choices, but not very big though.

4.4.2 Estimates from ordered logit

In above section, we got the results from simple logit by using the first best choice. To fully utilize the information and investigate whether the respondents make the choices consistently or not when they made the first and second choices, we next estimate the ordered logit model. To do this, we use the log likelihood function (18).Table 5 presents the results, and for convenience of comparison between simple logit and ordered logit, we also reproduce table 2 here.

Table 5. The ordered logit, specification (26); and simple logit, specification (26)

		Ordered logit model		Simple logit model	
Variable	Coef	Estimate	t-value	Estimate	t-value
Dummy for TA	β_{TA}	-0.5862	-1.8906	-0.5486	-1.2162
Dummy for CD	β_{CD}	0.1908	0.8684	0.0937	0.3137
Time until effect	β_t	-0.0585	-13.065	-0.0675	-11.9670
Lives saved	β_l	0.0084	3.0111	0.0091	2.4436
Dummy CD*life	β_{lCD}	-0.0056	-2.0128	-0.0059	-1.5926
Dummy TA*life	β_{lTA}	0.0069	1.8098	0.0034	0.6207
Cost (in 1000 NOK)	β_c	-0.1630	-8.3318	-0.1765	-6.4687
# of bs.			980		980
Log-likelihood			-2284.6		-1263.1
ρ^2			0.0618		0.0702
$\bar{\rho}^2$			0.059		0.0651

From Table 5, the estimates are quite similar in both simple logit and ordered logit models. And all the coefficients have the expected sign. That means when the respondents make the ranking choices, their behaviors are quite consistent.

The time and cost have the expected sign and sharply determined. Utility decreases in time and cost when holding other attributes constant.

The coefficients β_{lCD}, β_{lTA} from simple logit are not significant at any reasonable level of significance. Now β_{lCD} from ordered logit are significant at 5% LOS, although β_{lTA} is not significant. Recall in section 4.4.1, we also tested that at least one of β_{lCD}, β_{lTA} different from 0. In this sense the results from ordered logit are more reasonable.

The insignificance of β_{lTA} tells that the slope of life saved from TA can be considered the same as the life saved from EP. While the significance of β_{lCD} tells that the slope of life saved from CD is significantly less than that of life saved from EP.

It indicates that the utility for a respondent to save a life from environmental death is equal to the utility for the respondent to save 3 (=0.0084/(0.0084-0.0056)) lives from heart diseases, if all the other variables are held constant. And the respondent gains the same utility between saving a life from environmental pollution and saving 0.5490 lives from traffic accident.

4.4.3 VSL from ranking

According to Jones-Lee (1991, 1994), the value of statistical life is defined as the population mean of the marginal rate of substitution between wealth and risk. Then in our case, from the utility function specification, we can get:

$$VSL_i = POP \cdot \frac{\beta_i}{\beta_c}.$$

Where VSL_i is the estimate value of VSL for the death cause *i*. β_i and β_c are respectively the coefficients of the variables of number of lives (saved from death cause *i* (*i*=TA,CD,EP)) and the cost. Since in our data, the unit is household, we use the household number (given as 2 million) as POP here.

The VSL estimates in the following table:

Table 6. VSL estimates from ordered logit model, million NOK

	Heart diseases	Traffic accident	Environmental causes
Value of Statistical life	17.18	93.86	51.54

From Table 6, it's easy to see that the values of VSL from the ranking estimates are quite high.

This is not that surprising though. It agrees with the findings in Halvorsen and Sælensminde (1998). They claimed that the estimated WTP would be normally higher from the ranking estimates than from the CVM questions. They used some sophisticated technique nested logit to explain this. In next section, Here we’ll use a less technical and easier to understand approach to try to find out the reason.

5. Analysis of the dichotomous-choice

In the interview, after being asked the questions about the ranking problem, the interviewee is asked about whether he/she will actually like to pay for the cost of his/her most preferred project (denote it as *P*).

This answer can also be placed into a discrete choice framework. The choice set that the interviewee is facing is simply the project *P* and a project, which do nothing (denote it as *N*). The project *N* can be characterized with 0 value of all the attributes, namely save 0 life within 0 year and cost nothing. Then 'yes' answer to the dichotomous-choice question corresponds to the project P, while 'no' answer corresponds to the project N.

Proceeding as section 4, we write utility associate to those two projects as

$$U_P = v_P + \varepsilon_P$$

$$U_N = v_N + \varepsilon_N = \varepsilon_N$$

Since all the attributes associated with this project is 0, so the deterministic part of the utility function of saying no is simply 0, then we have the probability of answering 'yes' will be

$$P(yes) = P(\text{choose project P}) = \frac{e^{v_P}}{e^{v_N} + e^{v_P}} = \frac{e^{v_P}}{1 + e^{v_P}} \tag{27}$$

And the probability of answer 'no' will just be

$$P(no) = P(\text{choose project N}) = \frac{e^{v_N}}{e^{v_N} + e^{v_P}} = \frac{1}{1 + e^{v_P}}$$

So similar to section 4, using the utility specification (26), we are able to estimate the parameters of the utility function using the dichotomous-choice data.

If we think that the interviewees are fully rational, their preference should be coherent in the ranking procedure and the dichotomous choice situation --- The stated preference when it is the ranking questions being asked should not be different from the stated preference when the dichotomous-choice question being asked. So we would expect that the estimates of the same utility function should not be too far from each other. Of course, the estimates from the ranking problem should be more precisely estimated since there is more information available there.

In Table 7, we give the estimates based on the dichotomous choice, reproducing the estimates from the ranking questions for comparison.

Table 7. Estimates from the dichotomous-choice.

		Dichotomous-choice*		Ordered Logit	
Variable	Coef.	Estimate	T-value	Estimate	T-value
Dummy, causes TA	β_{TA}	-22.8370	-9.081	-0.5862	-1.890
Dummy, causes CD	β_{CD}	-14.9864	-17.382	0.1908	0.868
Time until effect	β_t	-0.0373	-1.673	-0.0585	-13.000
Number of lives	β_l	0.0630	3.264	0.0084	3.000
Dummy CD*life	β_{lCD}	-0.0698	-2.644	-0.0056	-2.000
Dummy TA*life	β_{lTA}	0.0959	4.359	0.0069	1.816
Cost	β_c	-18.8228	-10.101	-0.1630	-8.316

(* We had some problems to estimate this model under default convergence condition under TSP 4.5. The likelihood function is pretty flat near the optimal point. So we had sacrifice the precision a bit -- the tolerance is changed from 0.01 to 0.05)

We see immediately that some of the estimates are far from each other. The biggest difference occurs for the estimates of cost. This finding agrees with the conclusion from Halvorsen (2000) that respondents react differently to a discrete-choice CVM (DC-CVM) question than to a ranking one.

And it is very obvious that when confronted the dichotomous-choice question, the respondents put more emphasis on cost than in the ranking questions.

What could be the reason? We think that Halvorsen (2000) provided excellent explanations and discussions on this issue. From the following we will briefly reproduce her basic arguments. She pointed out that it might be caused by the fact that many respondents find both the ranking and DC questions quite difficult to answer. It is a sound explanation and the data from the survey support this view as well. First of all, in the questionnaire, the respondents are asked about how often they felt they could answer the four-project ranking questions according to their preferences. 28.8% answers ‘no’ to this question, i.e., almost one third of the respondents have problem to understand the questions. And when asked whether they put any priority in the attributes when they made the choice for the best and second choices, 69% of the respondents stated that they put focus on the number of lives saved when they answered the ranking questions. Second, we also feel that the way that those questions are asked plays a role here as well. Different ways of presenting the same question may trigger the respondents to put emphasis on different aspects of the projects. And thirdly, there are also possibilities that when asked about the dichotomous-choice and the Willingness-to-pay questions, the respondents may systematically understate their true willingness-to-pay for some kind of strategic concerns.

The above fact shows that DC method is still not yet a perfect substitute for open ended contingent valuation in practice, despite the fact that the NOAA (The National Oceanic and Atmospheric Administration) panel on contingent valuation suggested that discrete choice contingent valuation format is preferred to open ended format. In an interview situation, the respondents are asked to rank some projects they are not very familiar with and supposed to give answers within a quite short time, it is not surprising that they will have to make some kind of simplification. We believe that how to develop a more efficient way to conduct contingent valuation is still an open question, which needs more economists, maybe psychologists as well to work on.

6. Willingness-to-Pay and value of statistical life

In the above analysis, by comparing the results from the ranking procedure and the dichotomous-choice, we showed that it is not a very consistent way to derive VSL from the ranking questions. So in this section, we will try to make use of the stated WTP value for the first best choice to calculate VSL instead. We use a simple WTP regression first.

In contingent valuation studies, the value of a statistical life is equal to the average willingness to pay divided by the reduced risk of death (dR). In this case, the reduced risk of death is (in general, the reduced risk of death is equal to the number of lives saved divided by the affected population). That is, $VSL = \frac{\overline{WTP}}{dR}$, where $\overline{WTP}$ is average WTP. For example, if the average WTP = NOK 500 and dR = .0001 (1 in 10,000), then VSL = 500/. 0001 = NOK 5 million.

Above case is the simplest case. In our case, besides the reduced risk of death, there are still some other attributes: the time lagged before the project becomes effect, the cost to carry out the project, and etc. So, it's not the pure relationship between WTP and the reduced risk of premature death. Since the WTP we get here is only related to the most preferred project of the respondent, there might exist selection problem.

6.1 WTP regression

We here do a regression of WTP on the project's attributes, namely the number of lives saved by the effect and time until effect of project. As we discussed in section 4, we suspect that people value the different death causes differently. So saving one more life due to traffic accident is not the same as saving one more life due to heart disease. We should have three interaction terms of number of life saved and dummy of death causes instead of just one life variable. It is obvious that we should not simply put them as linear additive term in there, since the WTP should be 0 when we have zero life saved in the project.

Thus, we choose to use the following specification

$$WTP_n = \tilde{\beta}_{\mathrm{TA}} D_n^{TA*} l_n^* + \tilde{\beta}_{\mathrm{CD}} D_n^{CD*} l_n^* + \tilde{\beta}_{\mathrm{EP}} D_n^{EP*} l_n^* + \tilde{\beta}_t t_n^* + \alpha_n l_n^* + \varepsilon^* \qquad (28)$$

Where $\alpha_n = \alpha_1 age_n + \alpha_2 gender_n + \alpha_3 income_n$. l_n^* is life saved in the first preferred project, t_n^* is time lag in the first preferred project, *D* is 0,1 dummy for death causes in the first preferred project.

By using this α_n , it means that the coefficients of number of lives saved depend on the socio-economic variables. Here we also assume that the coefficient differences between these three death causes is the constant for all the respondents, i.e. $\tilde{\beta}_{TA}, \tilde{\beta}_{CD}, \tilde{\beta}_{EP}$ are identical for all the respondents. Of course, this assumption about the coefficient differences is not necessary, and it might be more reasonable to let all the interaction coefficients depend on the socio-economic variables. But we won't try it in this paper. We might perform it for further study in future.

VSL then can be calculated as

$$VSL = \frac{\Delta WTP}{\text{risk reduction}} = \frac{\Delta WTP}{\Delta L / population} = \frac{dWTP}{dL} Population \qquad (29)$$

It is easy to see that VSL will be different for different persons, and even for the same person we will have one value for each death cause. So using this specification, we recovered the distribution of the VSL. This may be different from the one we usually see in the literature. But to reach the VSL in the common sense, we could simply calculate the mean. If we use stratified sampling instead of random sampling used here, we will be able to infer that the mean VSL of the whole population will simply be the weighted mean.

One simple example to justify the formula (29) is to consider the simplest case, where we ask everyone in our sample the same open-ended question about WTP for one project that saves a certain number of lives (denote as $L_n(or\ L)$). And we have no other socio-economic variables. Analog to (28), we have a very simple regression

$$WTP_n = aL_n + \varepsilon$$

And it is easy to see that the estimated value for a will just be (and recall that $L_n = L$)

$$\hat{a} = \frac{\sum_i WTP_i}{\sum_i L_i} = \frac{\overline{WTP}}{L}$$ Where $\overline{WTP}$ is the sample average of WTP

So according to formula (29),

$$VSL = \frac{\overline{WTP}}{L / Population}$$

Which is exactly the formula we are very familiar with (see beginning of section 6).

6.2 The probability of selection bias

Generally, there will be no definite relationship between the WTP for proposed projects and ranking among them, when costs of the projects are taken into account in the ranking procedure. So the fact that we only observe WTP for the most preferred projects doesn't really give rise to a self-selection problem. However, in our study, among 980 observations, only 8% of the respondents took the project cost as the sole decisive attribute. And 69% respondents put focus on the number of lives saved when they answered the ranking questions. That is, the respondents simplify the ranking problem by focusing on some particular attribute, while ignoring other attributes. From the above numbers, we suspect that cost here doesn't make an important role in the ranking procedure, although we did get a sharply determined coefficient from the discrete choice analysis in the previous section. So as a result, we suspect that there exists a monotonic relationship between WTP and ranking. The consequence of this problem has two aspects. It suggests that we will not be able to estimate consistent value for WTP directly from the estimated 'preference' – we will always overestimate the WTP from the 'stated preference'. This also gives an explanation for the overestimated WTP value from discrete choice model estimates as we have discussed in section 4. Since we are only able to observe the WTP value for the most preferred project, which in our case might be the project with highest WTP, there will be endogenous selection involved in our analysis. The simple regression procedure will suffer from the danger of selection bias.

To correct the selection bias, there is possibility to use 'Heckman two-step method'. But following we will try to develop a different method which is quite similar to 'Heckman two-step method'. This method corrects the selection problem that arises because of lacking of consideration of cost in the ranking procedure. Under the assumption that the cost of the project is not involved into the ranking procedure, we will link the WTP with the random utility framework we used in previous two sections, and find a new way to make use of the information of WTP.

Based on our assumptions, we can write

$$WTP_n = f_n(U_n^*) \tag{30}$$

In(30), WTP_n is the willingness to pay of respondent n for his/her first best choice; and U_n^* is the utility for respondent to choose his/her first best choice.

Or equivalently,

$$U_n^* = f_n^{-1}(WTP_n) \tag{31}$$

Here f is a strictly increasing function, and U is the random utility function we defined in the previous two sections.

So to begin with, we will first establish some results for U.

Since we can only observe the WTP for the first best choice, we will only have observations on U for the first best choice as well. It is obvious that the observations follow the conditional distribution $P(U_j \le x \mid j \text{ is the best choice})$. So the problem transforms into finding this conditional distribution.

Theorem: .(from Dagsvik (2000) p93 lemma 2)

Suppose $U_j = u_j + \varepsilon_j$, where $(\varepsilon_1, \varepsilon_2 \cdots \varepsilon_m)$ is multivariate extreme distributed. Then $\Pr(\max_k U_k \le y \mid U_j = \max_k U_k) = \Pr(U_j \le y \mid U_j = \max_k U_k) = \Pr(\max_k U_k \le y)$

Apply this theorem in our case, we will have

$\Pr(U_j \le y \mid j \text{ is the best choice}) = \Pr(U_j \le y \mid U_j = \max_k U_k) = \Pr(\max_k U_k \le y)$

So we now only need to find out the distribution for $U_n^* = \max_k U_{nk}$. Note $U_{nj} = u_{nj} + \varepsilon_{nj}$, so U_{nj} is extreme value distributed with location parameter u_{nj} and scale parameter σ. Since any monotonic transformation of the utility function is still utility function. So, we have

$$U_{nj} = u_{nj} + \varepsilon_{nj} \quad \Leftrightarrow \quad \sigma U_{nj} = \sigma u_{nj} + \varepsilon'_{nj} \quad \Leftrightarrow \quad U'_{nj} = v_{nj} + \varepsilon'_{nj}$$

Here $v_{nj} = \sigma u_{nj}$ and ε'_j is extreme value distributed with parameter (0,1). Then following from the property of the extreme value distribution (Ben-Akiva and Lerman (1985)), U_n^* is also extreme value distributed, and it has the location parameter $\ln(\sum_k \exp(v_{nk}))$ and scale parameter 1.

Thus, we can write

$$U_n^* = f^{-1}(WTP_n) = \ln(\sum_k \exp(v_{nk})) + \tilde{\varepsilon} \tag{32}$$

Where $\tilde{\varepsilon}$ is standard extreme value distributed.

We see it immediately that without any assumption about the functional form of *f*, there is nothing we can do. Theoretically, any strictly increasing function can do the job.

A very common and simple functional form used here is linear function, i.e.

$$f(x) = ax + b \quad a > 0 \tag{33}$$

Insert (32) into (33), we have

$$WTP_n = a\ln(\sum_k \exp(v_{nk})) + a\tilde{\varepsilon} + b \tag{34}$$

Here the right hand variables of (34) should include all the relevant alternative's attributes.

However, note that

$$
\begin{aligned}
WTP_n &= a\ln(\sum_k \exp(v_{nk})) + a\tilde{\varepsilon} + b \\
&= a\ln(\frac{\sum_k \exp(v_{nk})}{\exp(v_n^*)} \cdot \exp(v_n^*)) + a\tilde{\varepsilon} + b \\
&= a\ln(\frac{1}{p_n^*(i)}) + av_n^* + a\tilde{\varepsilon} + b
\end{aligned} \tag{35}
$$

In (35), $p_n^*(i)$ is the probability of choosing the first best choice for respondent *n*; and v_n^* is the corresponding deterministic part of the U_n^*.

With linear specification of the deterministic part of the utility function, for example, for respondent *n* to choose first most-preferred project,

$$u_n^* = X_n^* \beta \tag{36}$$

Where X_n^* the vector of the attributes of the first preferred project for respondent *n*,

Then,

$$v_n^* = \sigma X_n^* \beta \tag{37}$$

We will reach the linear regression form:

$$WTP_n = \tilde{a} - a\ln(P_n^*(i)) + X_n^* \vec{\beta}^* + \varepsilon^* \tag{38}$$

From the above calculation, we see immediately that in order to correct the specification error of omitting the attributes for the relevant alternatives, we need to include an 'instrument variable' $\ln(P_n^*(i))$ in the regression. The problem with this is that we are not able to observe $\ln(P_n^*(i))$. So we need to use the estimates from either the logit or the ordered logit model to predict it first.

6.3 Estimates from regressions

In this section, we give the estimate results of the regressions with and without selection. We are interested in how large the selection bias is, or in other words, how dangerous it is to ignore the selection problem. Of course the result is only based on our data set and we need to be more careful when we make any general claim on this issue.

Recall equation (38), we have

$WTP_n = \tilde{a} - a\ln(P_n^*(i)) + X_n^* \vec{\beta}^* + \varepsilon^*$,

Using regression specification (28), we have

$$WTP_n = \tilde{a} - a\ln(P_n^*(i)) + \tilde{\beta}_{\mathrm{TA}} D_{n,i}^{TA} l_{n,i} + \tilde{\beta}_{\mathrm{CD}} D_{n,i}^{CD} l_{n,i} + \tilde{\beta}_{\mathrm{EP}} D_{n,i}^{EP} l_{n,i} + \tilde{\beta}_t t_{n,i} + \alpha_n l_{n,i} + \varepsilon^*$$

Where $\alpha_n = \alpha_1 age_n + \alpha_2 gender_n + \alpha_3 income_n$

$l_{n,i}$ is life saved, $t_{n,i}$ is time lag, D is 0,1 dummy for death causes.

As we discussed in section 4.4.2, the results from ordered logit are more reasonable. So we will use the results from ordered logit to predict the probability for one respondent to choose his most-preferred project.

Now let us consider a situation as follows: If there is no project, that is, the value of all the attributes is 0, then, it's easy to see, our best choice is to choose maintain the original status, then $P_n^*(i) = P_n^*(\text{no payment}) = 1$. Then we have $WTP_n = \tilde{a}$. Obviously, we should set $\tilde{a} = 0$, since $WTP_n = 0$, when there is no project.

Now the linear regression becomes:

$$WTP_n = a\ln(P_n^*(i)) + \tilde{\beta}_{\mathrm{TA}} D_{n,i}^{TA} l_{n,i} + \tilde{\beta}_{\mathrm{CD}} D_{n,i}^{CD} l_{n,i} + \tilde{\beta}_{\mathrm{EP}} D_{n,i}^{EP} l_{n,i} + \tilde{\beta}_t t_{n,i} + \alpha_n l_{n,i} + \varepsilon^* \quad (39)$$

Running this regression, we get the results in Table 8.

Table 8. Impact on WTP (in 1000 NOK) of key background variables and design variables

		OLS without selection		OLS With selection	
Variable	Coef.	Estimate	t-statistic	Estimate	t-statistic
Time until effect	β_t	0.045435	4.96	-0.023121	-2.09
Dummy EP*life	β_{lEP}	0.014417	4.63	0.006423	2.09
Dummy CD*life	β_{lCD}	0.005877	4.30	0.003340	2.52
Dummy TA*life	β_{lTA}	0.022226	9.30	0.011007	4.35
Income*life	α_1	0.000004	1.38	0.000005	1.94
Age*life	α_2	-0.000042	-1.80	-0.000037	-1.67
Gender*life (woman=1)	α_3	-0.001028	-1.29	-0.000760	-1.00
-Log (pi)	*a*			1.072590	10.09

Note: Income is in 1000 NOK.

From Table 8, we notice that the coefficient of the 'instrument variable' log (1/pi) is significantly different from 0 and positive. And from the estimates of other coefficients, we find big changes. Note that the sign of β_t is positive in the regression without selection, which means the WTP increases in the time lagged, obviously this cannot be true. While in the regression with selection, β_t now has the expected sign. And the coefficient associated with age, gender and income doesn't change much. Most important, coefficients in front of the death causes dummy variables changed pretty much. So in our study, there is a sign of some selection biases.

And the dummy for CD causes is significantly determined at reasonable level of significance. It also indicates that 'age', 'number of cars' and 'number of kids' don't matter much. And gender might matter about the decision. Now, since the coefficient associated with gender is negative (Recall 1 refers to woman), it means man is willing to pay more for the chosen project than woman does, if all the other variables are held constant.

To calculate VSL, recall formula (29),

$$VSL = \frac{\Delta WTP}{\text{risk reduction}} = \frac{\Delta WTP}{\Delta L / population} = \frac{dWTP}{dL} Population$$

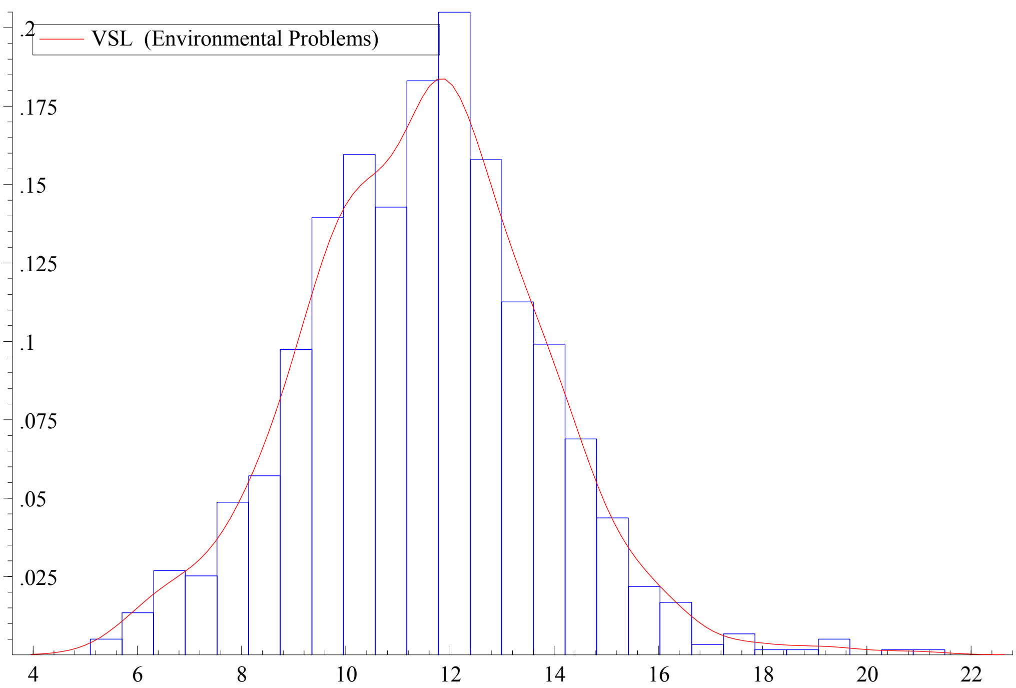
Given our specification of the regression, we have that

$VSL_n(cause) = \tilde{\beta}_{\text{Cause}} + \alpha_1 age_n + \alpha_2 gender_n + \alpha_3 income$.

Where *cause=TA (Traffic accident), EP (environmental problem), CD (heart disease)*

So we can calculate the VSL for every individual for all three causes. The following figures shows that the detailed VSL distribution in our sample for environmental problem. The other two VSL distributions are just simple shift along the X-axis, and the shape will be exactly the same. This is the consequence of our assumption, of course we could let α depends on those death cause dummies as well, then the distribution will have different shapes.

Figure 1. The distribution of VSL (Environmental Problems)



And the average VSL in the sample based on the regression estimation (with selection) for each cause is given in Table 9.

Table 9. Average VSL estimates based on the regression estimation, million NOK.

Model applied	Environmental causes	Traffic accidents	Heart diseases
OLS with selection	11.48	20.64	5.31

Comparing the VSL estimates in Table 9 and Table 6, we can see that the VSL calculated from the ranking preferences is much bigger than the one from open-ended questions. This is exactly what is found out in Halvorsen and Sælensminde (1998). But to look at the values in Table 9, VSL is suspiciously low, especially VSL related to heart diseases. We might suspect that the correction of the selection bias might have some negative effect on the estimation of VSL of the heart disease. To investigate about the sample we use, there are 60% of the projects are of heart disease, which might explain there is not much selection bias in the estimates of VSL of heart disease. There are some possibilities to do some more research on this in future.

7. Conclusions

In this paper, first, we use discrete choice technique and estimate a simple logit and ordered logit model to recover the preference associated with ranking and the preference associated with the risk reduction by using the ranking data. The results from these two models are quite similar. We find out that people significantly prefer to reduce the premature death related to the environmental pollution than to reduce the premature death caused by heart disease. But no significant evidence indicates saving lives from environmental pollution is more preferred than saving lives from traffic accident, or vice versa.

We also calculated the VSL directly from preferences based on our estimates. But the VSL found here is surprisingly high, compared with other studies. This agrees with the findings in Halvorsen and Sælensminde (1998). They claimed that individuals react differently to a dichotomous-choice CVM question than to a ranking one. In her paper, Halvorsen (2000) used much more sophisticated technique nested logit. Here we place the dichotomous-choice answers into a simple ranking framework instead, and use an approach that is less technical and easier to understand to elicit that the ranking and the dichotomous-choice are not consistent.

Furthermore, we try to link the WTP with the random utility framework in this paper. We suggest a new way to make use of the information of WTP. By introducing an "instrument variable" $z = \ln(P_n^*(i))$ into the regression, we can succeed to correct the selection bias.

We show that in theory the common estimates on study of the relationship between WTP and other socio-economic variables by using OLS is biased due to the selection problem. And preliminary study shows that danger of ignoring the selection problem does exist when we compare the empirical results from these two methods.

Reference:

Afriat, S. (1972).'Efficiency Estimation of Production Functions', International *Economic Review* 13, 568-598.

Adamowicz, W., J. Louviere and M. Williams (1994). 'Combining revealed and stated preference methods for valuing environmental amenities'. *Journal of Environmental Economics and Management* 26, 271-292.

Alpizar, F., Carlsson, F.,Martinssona, P.(June 2001). 'Using Choice Experiments for Non-Market Valuation' *Working Papers in Economics no. 52*,Department of Economics, Göteborg University.

Ben-Akiva, M. and S.R. Lerman (1985). 'Discrete Choice Analysis, Theory and Application to Travel Demand', The MIT Press, Cambridge, Massachusetts, London, England.

Berndt.E.K., B.H.Hall, R.E.Hall, and J.A.Hausman (1974). 'Estimation and Inference in Nonlinear Structural Models', *Annals of Economic and Social Measurement*, October 1974, pp. 653-665.

Blomquist, G.C., 1979. 'Value of Life Saving: Implication of Consumption Activity'. *Journal of Political Economy* 87, 540-558.

Boxall, P., W. Adamowicz, J. Swait, M. Williams and J. Louviere (1996). 'A comparison of stated preference methods for environmental valuation'. *Ecological Economics* 18, 243-253.

Carlsson, F. and P. Martinsson (2001). 'Do hypothetical and actual marginal willingness to pay differ in choice Experiments? – Application to the valuation of the environment'. *Journal of Environmental Economics and Management* 41, 179-192.

Blamey, R., J. Bennett, J. Louviere, M. Morrison and J. Rolfe (2000). 'A test of policy labels in environmental choice modeling studies'. *Ecological Economics* 32, 269-286.

Ciriacy-Wantrup, S. (1947). 'Capital returns from soil-conservation practices'. *Journal of Farm* Economics 29, 1181-1196.

Dagsvik, John K. (2002). 'Probabilistic Models for Qualitative Choice Behavior: An introduction'. *Statistics Norway Documents* 2000/1.

Davis, R. (1963). 'The Value of Outdoor Recreation: An Economic Study of the Maine Woods'. *Ph.D. dissertation*, Department of Economics, Harvard University.

Diewert, W.E. and C. Parkan (1983). ‘Linear Programming Tests of Regularity Conditions for Production Frontiers’, in *Quantitative Studies on Production and Prices* (W. Eichhorn, R. Henn, K. Neumann and R.W. Shephard, Eds.), Physica-Verlag, Würzburg.

Drèze, J. (1962).‘L’Utilitè Sociale d’une Vie Humaine,’ *Revue Française de Recherche Opèrationelle* 6: 93-118, 1962.

Garbacz,Christopher. 'Smoke Detector Effectiveness and the value of saving a life', *Economics letters*, Dec. 1989, 31(3), pp.281-286.

Ghosh, Debapriya; Lees, Dennis and Seal, William. 'Optimal Motorway Speed and Some Valuation of Time and Life', *Manchester School of Economics Soc. Stud.*June 1975, 43(2), pp.134-43.

Greene, W.H. Econometric analysis, fourth edition, New York University.

Halvorsen, B. (2000). ‘Comparing ranking and contingent valuation for valuing human lives, applying nested and non-nested logit models’. *Environment and resource economics.*

Halvorsen, B. , Strand, J. , Sælensminde, K. and Wenstøp, F. (1996). ‘Comparing Contingent Valuation, Conjoint Analysis and Decision Panels: An Application to the Valuation of reduced Damages from Air Pollution in Norway’. *Memorandum* from Department of Economics, University of Oslo.

Halvorsen, B. and Sælensminde, K. (1998).‘Differences between willingness-to-pay estimates from open-ended and discrete-choice contingent valuation method: the effect of heteroscedasticity’. *Land economics*, 74 262-282.

Garrod, G. and K. Willis (1998). ‘Estimating Lost Amenity due to Landfill Waste Disposal’, *Resources, Conservation and Recycling*, 22(1-2), 83-95.

Hanley, N., R. Wright and W. Adamowicz (1998). ‘Using choice experiments to value the environment’. *Environmental and Resource Economics* 11, 413-428.

Hanley, N., Shogren, J.F., White, B. (1997). ‘Environmental economics in theory and practice’. New York, Oxford: Oxford university press.

Hanoch, G. and M. Rothschild (1972). ‘Testing Assumptions of Production Theory: A Nonparametric Approach’, *Journal of Political Economy* 80, 256-275.

Jones-Lee, Michael W. (1994). ‘Safety and the Saving of Life: The Economics of Safety and Physical Risk’. *In Layard, R. and Glaister, S. (eds.) Cost-benefit analysis. 2nd edition.* Cambridge: Cambridge University Press.

Jones-Lee, Michael W. (1991). ‘Altruism and the Value of Other People’s Safety’, *Journal of Risk and Uncertainty* 4, 213-219.

Krupnick, A. J. et.al. (2002). ‘Age, health and willingness to pay for mortality risk reductions: A contingent valuation survey of Ontario residents’. *Journal of Risk and Uncertainty*, 24.

Kuriyama K., Takeuchi K., Kishimoto A. and Seo K. (May, 2002). ‘A Choice Experiment Model for the Perception of Environmental Risk: A Joint Estimation using Stated Preference and Probability Data’. *Presentation at the Second World Congress of Environmental and Resource Economists*, Monterey, California. P3.

Layton, D. and G. Brown (2000). ‘Heterogeneous preferences regarding global climate change’. *Review of Economics and Statistics* 82, 616-624.

Louviere, J., D. Hensher and J. Swait (2000) Stated Choice Methods. Analysis and Application. Cambridge: Cambridge University Press.

Mitchel R.C., Carson R.T (1989): ‘Using surveys to value public goods: The Contingent Valuation method’, Resources of the Future, Washington D.C.

Morrall, J. F. (1986). ‘A review of the record’. *Regulation*, 10, 25-34.

Norwood G., Maslow's Hierarchy of Needs. *http://www.connect.net/georgen/maslow.htm*, June 1996.

Portney Paul R. 'housing prices, Health Effects, and Valuing Reductions in Risk of Death', *Journal of Environmental economics Management*, March 1981,8(1).

Ryan, M. and J. Hughes (1997). ‘Using Conjoint Analysis to Assess Women’s Preferences for Miscarriage Management’. *Health Economics* 6, 261-273.

Rolfe, J., J. Bennett and J. Louviere (2000) ‘Choice Modelling and Its Potential Application to Tropical Rainforest Preservation’, *Ecological Economics*, 35(2), 289–302.

Rosen Sherwin.'The Theory of equalizing Differences' in Handbook of labor economics. Eds.: Orley Ashenfelter and Richard Layard. Amsterdam: North-Holland, 1986,641-692.

Schelling, T.C. (1968). 'The Life You Save May Be Your Own', *in S.B. Chase (ed.), Problems in Public Expenditure Analysis*, Brookings, Washington, D.C.

Strand J. (2002). 'Public- and Private- good values of statistical life'. *Working paper 2002: 2* Frisch Centre; Department of Economics, University of Oslo.

Statistics Norway (2002). 'Statistical Yearbook of Norway 2002'.

Tengs, T. O., Adams, M. E., Peskin, J. S., Safran, D. G., Siegel, J. E., Weistein, M. C. and Graham, J. D. (1995), 'Five-hundred life-saving interventions and their cost-effectiveness'. *Risk Analysis*, 15, 369-390.

Varian, H.R. (1984): 'The Non-Parametric Approach to Production Analysis', *Econometrica* 52, 279-297.

Varian, H.R. (1985): 'Non-Parametric Tests of Optimizing Behavior with Measurement Error', *Journal of Econometrics* 30, 445-458.

Varian, H.R. (1990), 'Goodness-of-Fit in Optimizing Models', *Journal of Econometrics* 46, 125-140.

Viscusi, W. K. (1993). 'The Value of Risk to Life and Health'. *Journal of Economics literature*, 31, 1912-1946.