Environmental and Resource Economics **8:** 485–499, 1996.

Ordering Effects in Contingent Valuation Surveys

Willingness to Pay for Reduced Health Damage from Air Pollution

BENTE HALVORSEN
Centre for Research in Economics and Business Administration, Gaustadaléen 21, N-0371 Oslo, Norway

Accepted 20 February 1996

Abstract. This paper focuses on ordering effects in CVM surveys; how the expressed value of a particular good valued in a sequence of several goods depends on where in the sequence the good is valued. We use data from a Norwegian CVM survey focusing on WTP for a 50% reduction in air pollution from car traffic to test for the existence of ordering effects and to apply a test for internal consistency. We found considerable and significant ordering effects in our data, but were not able to reject the hypothesis of internal consistency. Based on our survey, we argue that ordering effects may be a result of rational choice. These effects are problematic if a sequential valuation procedure is applied to a simultaneous problem, and/or the respondents are given imperfect information about the decision problem.

Key words: contingent valuation, ordering effects, air pollution, health damage

1. Introduction

Most people regard good health as very important in their lives. Several international studies show that the quality of the environment may influence the population's health condition (for example, see Perry et al. 1982, and Goldsmith and Friberg 1977). A recent Norwegian dose-response study (Clench-Aas et al. 1989) shows that changes in emissions from car traffic have a significant effect on the probability of contracting lung disease, such as chronic bronchitis, and several minor health problems. It is often possible for the government to take action to reduce these negative external effects. In order to decide whether or not to take action, the government may want some measures of the population's preferences for these actions. For this purpose the Contingent Valuation Method (CVM) is often used.

A project to reduce emission from car traffic will, in addition to the effects on human health, also have effects on agricultural and forestry production, and damages to the natural environment due to acid rain. Problems may occur in the valuation of the project if not all effects are described in the questionnaire, or if we want to isolate the value of a particular effect, for instance the effect on human health. Hoehn and Randall (1989) and Randall (1991) discuss aggregation biases that occur if a project is valuated on its own when it is a part of a more general policy agenda.[1] They show that the sum of the partially valued projects on the agenda[2] exceeds the simultaneous valuation[3] as the number of projects on the agenda increases. This is due to income and substitution effects in a general equilibrium framework. In their concluding remarks, Randall and Hoehn say that the only way to obtain the total value of all projects on a policy agenda is: '*... either a one-shot, holistic valuation or a sequenced approach*' (p. 550). There are problems, however, with both these valuation procedures. The one-shot valuation of the whole agenda does not value individual projects. This may be done with a sequential procedure, but then the value of a particular project is not unique (Hoehn and Randall 1989).

Over the last few years, a discussion has taken place questioning the Contingent Valuation Method's ability to value a particular good in a multidimensional valuation problem. In a critical article, Kahneman and Knetsch (1992) emphasize two different groups of effects: first, *ordering effects*, where the value of a particular good as perceived by the respondents depends on where in a sequence it is valued; when a given set of goods are valued in a sequence. Second *embedding effects*, where the value of a particular good as perceived by the respondents is sensitive to the number of goods to be valued. Another critique of CVM is put forward by Diamond and Hausman (1994). They are, among other things, concerned about the lack of testing for internal consistency in the valuation literature, as these tests are important to assess the reliability and validity of CVM surveys. They recommend an 'adding-up' test to test whether the sums of a valuation sequence equal the direct total value.[4] That is to test whether Hoehn and Randall (1989) are correct when they claim that the sequential valuation method will estimate the total value of the agenda correctly. Diamond and Hausman claim that: '*When these tests have been done, contingent valuation has come up short*' (p. 62), referring to Samples and Hollyer (1990).

In their article, Samples and Hollyer (1990) apply this adding-up test, and reject the hypothesis of internal consistency. This test is also applied in a study by Hoevenagel (1994) with the same results. Both Samples and Hollyer (1990) and Hoevenagel (1994) also test for the effect of sequencing on the willingness to pay (WTP) for a particular good, which is what Kahneman and Knetsch (1992) call the 'ordering effect'. Both found significant effects.

In this paper, we use data from a CVM survey to value the Norwegian public's willingness to pay for a governmental program reducing the emissions to air from car traffic by 50% to test for ordering effects and internal consistency. The results from our survey contradict the critique put forward by Diamond and Hausman (1994) by, in contrast to the results from Samples and Hollyer (1990) and Hoevenagel (1994), failing to reject the hypothesis of the adding-up test. This conclusion holds in every analysis we conducted, and we therefore believe it to be a very robust result. We also find considerable and significant ordering effects. Our results imply that these ordering effects are made by respondents that are, on average, internally consistent. We thus argue that ordering effects may be the result of a rational choice, and that they are unwanted if the respondents are not provided with full information about the decision problem at the decision point.

The rest of this paper is organized as follows: section 2 contains a description of the economic theory behind the survey. Here we model the WTP, discuss how to decompose the total WTP, and present our hypothesis. Section 3 describes the questionnaire and the survey. In section 4 we present the results of the analysis, testing the hypotheses about ordering effects and internal consistency. This is done by applying both descriptive statistics and various regression analyses. In the final section, some concluding remarks are made considering the effect of imperfect information on the results from a CVM survey.

2. The Theoretic Framework for the Survey

2.1. Modelling Willingness to Pay

In the model, we consider a representative consumer whose two-period utility depends on the consumption of private goods (X in period one, and Z in period two), his health condition (h), and on damages to the natural environment (K). The consumer's expected utility depends on whether he will be ill (s) or well (w) in the future, with probabilities π and $1-\pi$, respectively. The individual's probability of becoming ill (π) and the damage to the natural environment (K) both depend on the air quality; an improvement in air quality will lower the probability of becoming ill, and reduce damage to the natural environment, *ceteris paribus*. We assume that an improvement in air quality in period one will reduce the probability of becoming ill in the next period only. However, damage to the natural environment is affected in both periods. We assume that the consumer minimizes his total expenditures over the two periods subject to a given total discounted expected utility.[5] If the government does not take action to improve air quality, the consumer will not take this factor into consideration, because air quality is a public good. The total expenditure function is then defined by:

$$C = C(p_x, p_z, \bar{U}, h_s, h_w, h, K, \pi) = \min_{X,Z} p_x X + p_z Z$$

$$\text{s.t. } [\pi U(Z; h_s, K) + (1-\pi) U(Z; h_w, K)] + U(X; h, K) = \bar{U}. \quad (1)$$

Suppose the government can guarantee a 50% improvement of the air quality. This improvement will benefit the consumer because it will reduce both π and K. The price the individual has to pay for a reduction in air pollution is the increase in tax revenue necessary to finance governmental actions. The individual's compensating variation (CV) for an increase in air quality is the maximum amount of income he is willing to forgo to be indifferent to no improvement at all. We will denote the initial state (π^0, K^0), and the state after an improvement in air quality (π^1, K^1) where $\pi^1 < \pi^0$, and $K^1 < K^0$. The consumer's willingness to pay for this air quality improvement is given by:

$$\text{CV} = C(p_x, p_z, h, \bar{U}, \pi^0, K^0) - C(p_x, p_z, h, \bar{U}, \pi^1, K^1). \quad (2)$$

This CV measure will be positive if air quality is improved, and negative if it is reduced. The reason is that the consumer needs to use less income to maintain a given utility level when air quality improves.

2.2. Decomposition of Total Willingness to Pay

In our model, the WTP is motivated either from a desire to reduce the risk of becoming ill or to reduce the damages to the natural environment. The government may want to decompose the total WTP into WTP for the different motives to learn more about the people's preferences regarding the project. A decomposition of the total of WTP is also necessary to test for internal consistency and ordering effects.

This decomposition can be done in two different ways: the bottom-up, or the top-down approach. The bottom-up approach is also known as the sequential valuation method, and was suggested by Randall and Hoehn (1989) and Randall (1991). The WTP procedure is composed of separate questions, in our case two, which are asked in sequence: first the respondents are asked to state their WTP for a reduction in the risk of becoming ill in the future, after being given information about the health effects only. Subsequently, they are given information about damage to the natural environment, and asked to value the reduction in the damage to the natural environment, conditioned on their stated WTP for a reduction in the risk of becoming ill. In this sequential decision problem, the value of a good will depend on where in the sequence it is valued. The total WTP is divided into WTP for a reduction in the risk of becoming ill, conditional on no changes in K, $\mathrm{CV}(\pi|K = K^0)$, and WTP for reduced damage to the natural environment, conditional on the stated WTP for a reduction in π, $\mathrm{CV}(K|\mathrm{CV}\ (\pi|K = K^0))$. The second decomposition method is described as the top-down approach and implies a simultaneous decomposition of the total WTP into the two different benefit groups. Here the respondents are asked to value all effects of a project in a one-shot, holistic valuation. Then they are asked to distribute their total stated WTP into WTP for a reduction in the risk of becoming ill ($\mathrm{CV}(\pi)$), and the WTP for a reduction in damages to the natural environment ($\mathrm{CV}(K)$). This is done by weighting the importance of the two categories.

2.3. Comparing the Two Different Decomposition Methods

The sequential valuation of a reduction in the risk of becoming ill is likely to be greater than, or equal to, the simultaneous weighted WTP. The opposite is the case for the WTP for reduced damages to the natural environment. This is due to the income and substitution effects. When a consumer with a given income is introduced to a new good (that is when the feasible set of goods previously endogenous to the decision problem increases), and the consumer derives utility from consumption of this new good, the proportion of the total expenditure on all other goods declines as the number of available goods increases (Randall and Hoehn 1989). If the respondents are asked to value changes in the probability of becoming ill alone, the WTP for these changes will be higher (as a proportion of gross income) than if changes in the probability of becoming ill *and* the damages to the natural environment are valued simultaneously.[6] Thus, the decision problem in the survey should not endogenize more (or less) goods previously exogenous to the decision maker than those relevant to the decision problem. If it does, the questionnaire is likely to create what Kahneman and Knetsch call embedding effects. In addition, imperfect information about the feasible set of goods in the decision problem, as the respondents are given additional information during the valuation sequence, may leave a perfectly rational respondent with an inoptimal allocation of income *ex post*. Both these effects will make the sequential WTP for a reduction in π exceed the simultaneous decomposed WTP, creating ordering effects. If the goods are valued in a sequence when they ought to be valued simultaneously, for example when a project affects more than one good, this may result in biased estimates of the WTP for these goods. We will call this bias unwanted ordering effects. The existence of these ordering effects does not necessarily imply that the respondents are irrational or boundedly rational.[7] It merely reflects the information problems they are facing at the decision point. However, if the respondents are allowed to reconsider their valuation after all information is received, the added total value should be equal to the direct total value, and the hypothesis of internal consistency ought to hold.

This imperfect information problem may have been one reason why Samples and Hollyer (1990) and Hoevenagel (1994) had to reject the hypothesis of internal consistency. In their survey, Samples and Hollyer valued a project saving both seals and whales. They found that the sum of the sequential valuation of the project depended on whether the whales were valued before or after the seals. They assumed that the reason for this difference was that whales are generally more popular than seals, and the respondents wanted this to be reflected in their valuation. When the whales were valued after the seals, the WTP for the whales was anchored on the valuation of the seals. Hoevenagel (1994) valued three different programs: (i) restoring Dutch woodlands, (ii) reducing noise from air and road traffic, and (iii) reducing the greenhouse effect. Hoevenagel had to reject both the hypotheses that sequencing does not affect the total WTP, and that 'the direct total value' equals 'the added total value'. The information problems in Hevenagel's survey may have been more severe than in Samples and Hollyer's, since the respondents were neither allowed to reconsider their initially stated WTP after all information was received, nor were they explained that all contributions made for the second and the third good would be in addition to the payments made for the first good (Hoevenagel 1994: 171). These information problems make it difficult to draw strong conclusions from the rejection of the hypothesis. In our survey, we tried to avoid the problems imperfect information causes on the tests for internal consistency by asking the respondents to state a new total WTP for all goods when all information is revealed.

2.4. THE HYPOTHESES

Our first hypothesis is concerned with the internal consistency of the WTP statements. According to Randall (1991), the one-short, holistic valuation of all effects should be equal to the sum of the sequential valuation of each good (see Randall 1991: 306, equations (10.4) and (10.5)). This is basically the adding-up test suggested by Diamond and Hausman (1992). Thus, we do not expect the sum of the valuation sequence to be significantly different from the direct total WTP.

(H-1) H_0: $\text{CV}(\pi|K = K^0) + \text{CV}(K|\text{CV}(\pi\ K = K^0)) = \text{CV}(\pi, K)$

H_1: $\text{CV}(\pi|K = K^0) + \text{CV}(K|\text{CV}(\pi\ K = K^0)) \neq \text{CV}(\pi, K)$.

The next two hypotheses concern the unwanted ordering effects discussed in section 2.3. Due to the income and substitution effects and the imperfect information about the feasible set of goods, we would expect a sequential valuation procedure to overstate the true value of the first and understate the true value of the second good in the valuation sequence.

(H-2) H_0: $\text{CV}(\pi|K = K^0) = \text{CV}(\pi)$

H_1: $\text{CV}(\pi|K = K^0 > \text{CV}(\pi)$

(H-3) H_0: $\text{CV}(K|\text{CV}(\pi|K = K^0)) = \text{CV}(K)$

H_1: $\text{CV}(K|\text{CV}(\pi|K= K^0)) < \text{CV}(K)$.

3. Design of the Survey

The survey was part of a nationwide monthly Omnibus in May 1993 with 1229 personal interviews. The respondents were given a scenario where the benefits from a 50% reduction in air pollution due to reduced emissions from traffic were described. The benefits mentioned were (i) a reduction in the risk of becoming ill from lung disease, asthma, bronchitis, allergy, and minor health effects such as a reduction in days with headache, tiredness, aching muscles, cold or flu, and (ii) a reduction in damage due to acid rain: damage to forestry and agricultural production, and material damage.

As an example, we provided the respondents with a description of possible health effects due to a 50% reduction in air pollution. The estimated health effects were based on the results from a dose-response survey done in one of the most heavily polluted parts of Oslo (Clench-Aas et al. 1989). The respondents were told that the government would take the necessary action to implement the desired improvement in air quality, and that the increase in governmental costs would be financed through higher income taxes. To make sure that the respondents would consider their budget constraints, we explained how an increase in income taxes would reduce their purchasing power, and urged them to consider the limitations of their income before answering. The respondents were then asked to state the maximum sum of money they would be willing to pay in increased income taxes for this project.

We divided the sample into fur sub-samples with two splits. (i) Respondents in sub-samples B and D were told that the government would subsidize electric cars to achieve the 50% reduction in air pollution. The respondents in sub-samples A and C were told that the government would use an unspecified package of tools to achieve the required reduction in emissions from car traffic. (2) Respondents in sub-samples A and B were given all the information, and then asked to value all the benefits from a 50% reduction in the air pollution from cars. The respondents in sub-samples C and D were first given information about health effects and then asked to value these effects. Subsequently, they were told about all other effects, and asked to state a new total value for all the benefits mentioned. In this way, we made all the respondents reconsider their total WTP after all the information was given. The first split was made to reduce the conditioning of the WTP statements due to the suggested policy instrument. The second split made it possible to decompose total WTP by applying the bottom-up approach, and to test for ordering effects and internal consistency. The top-down approach was applied on all respondents, asking them to value all the benefits (in either one or two WTP questions) and then to distribute their total WTP according to different motives. The different motives mentioned were: (i) a reduction in the risk of becoming ill for the respondent and his closest family, (ii) a reduction in other people's risk of becoming ill, and (iii) reduced damage to the natural environment, production, and materials due to a reduction in acid rain. One potential problem with this method is that many people may have problems weighting the different motives. This may result in more uncertain answers, but probably not in systematic bias.

4. Analysis of the Results From the Survey

The main aim of this survey was to estimate the unwanted ordering effects and to test for internal consistency. This was done both by applying descriptive statistics and various regression models. Before commenting on these results, we will briefly comment on how we treated respondents that either stated a very high WTP as a proportion of reported annual gross household income, or those who stated a zero bid as a protest to the payment vehicle.

4.1. Protest and Very High Bids

Fifty eight per cent of the respondents stated a zero bid. Most of the zero bidders gave one of the following reasons for their answer:

Table I. The share (in percent) of the respondents with certain characteristics in the sub-samples of protest bidders and for all respondents.

Sample / Variable	The whole sample	CV = 0 due to too high a taxation level	CV = 0 but positive if voluntary contributions	CV = 0 but positive if the the government redistributes current revenue
Major city	22	18	10	17
University degree	21	10	17	20
Minor health problems	61	55	47	53
Chronic diseases	51	45	37	47
Gross houshold income	272 000	244 000	270 000	263 000

- 'I think the taxation level is high enough as it is' (48%);
- 'I am generally positive, but think the government should redistribute its current revenue' (40%), or
- 'I would want to contribute if the payment vehicle was organized in an other manner, for instance by voluntary contributions' (4%).

These respondents reacted negatively to the payment vehicle, and may be considered as protest bidders. However, the respondents were asked to value a project consisting of a 50% reduction in emissions, a policy instruments and a payment vehicle. To identify the properties of the marginal distribution for the environment effects only, we need to vary both the policy instrument and the payment vehicle to eliminate the conditioning they cause on the WTP statements. If we exclude the protest bidders, we would cause additional conditioning to the problem as these respondents are not randomly distributed among the sub-samples (see Table I). In our study, we varied the policy instruments but not the payment vehicle. Thus, we decided to include the protest bidders and interpret our results as the WTP conditional on the payment vehicle. However, excluding the protest bidders do not alter the conclusions in the paper.

In our sample, 60 out of the 1148 respondents who answered both the valuation question and the question about annual gross household income had a WTP exceeding their reported income. All these respondents reported a zero income. After some investigation we decided not to exclude any of these respondents for two main reasons: the first is that annual gross income is only a proxy for the relevant income variable in this problem, which is the permanent income. Twenty-six out of these 60 respondents were students that had not yet entered the labour market, and two were temporarily out of employment. Second, the income variable is obviously not correct for some of these respondents. Twenty out of the 60 respondents who reported a zero household income also said that they had a steady job, but for some reason they did not want to report their income. Three were retired and were thus receiving a pension from the state. From this, it is obvious that the income variable is problematic. Thus, we cannot conclude that there is something

Table II. Mean decomposed WTP (NOK) for a reduction in the risk of becoming ill and damage to the natural environment, and test statistics for the hypotesis (H-2) and (H-3).

	Mean	t-value
Sequential valuation of a reduction in the risk (π)	1133	
Simultaneous valuation of a reduction in the risk (π)[a]	278	$T_{H\text{-}2} = 9{,}6$
Sequential valuation of a reduction in damage (K)[b]	6	
Simultaneous valuation of a reduction in damage (K)[a]	862	$T_{H\text{-}3} = -17{,}2$

[a] These values are derived by multiplying the total WTP with the weight placed on the different benefits.
[b] This value is derived as the residual between the total WTP and the sequential decomposed WTP for a reduction in π for the respondents receiving two valuation questions.

wrong with the WTP variable from the fact that some respondents have a WTP exceeding their household income.

4.2. Descriptive Statistics

We found in section 2.3 that the simultaneous decomposed WTP for a reduction in the risk of becoming ill should be less than or equal to the sequential decomposed WTP, and vice versa for the other benefits. The mean willingness to pay for the different motives, and the test statistic for the two hypotheses (H-2) and (H-3) are given in Table II. The mean simultaneous decomposed WTP for a reduction in the risk of becoming ill, CV(π_s), was 278 NOK, and the sequential decomposed WTP, CV($\pi | K = K^0$), had a mean of 1133 NOK, which is about four times higher. For the rest of the benefits mentioned, the simultaneous decomposed WTP exceeds the sequential, as expected in the hypothesis (H-3). We reject the hypothesis of no ordering effects as the test statistics for the two hypotheses (H-2) and (H-3) are clearly significant ($T_{H\text{-}2} = 9.6$ and $T_{H\text{-}3} = -17.2$). These differences may also be due to part-whole biases which will occur if the respondents are aware of the other effects, but are given limited information about the valuation procedure. This implies that a sequential valuation is likely to yield biased estimates for the value of a particular good. One reason for these results is that imperfect information at the decision point makes it difficult for the respondents to adjust optimally to the multidimensional decision problem.

In Table III, we present the descriptive statistics for the second split; whether the respondent answered one or two valuation questions. The mean WTP in the two-samples are approximately the same, and the t-value for the hypothesis of internal consistency is very small ($T_{H\text{-}1} = 0.04$). This implies that whether the respondent answered one or two WTP questions had no significant effect on the total WTP. Since we are not able to reject the hypothesis (H-1) it suggests that the respondents, on average, behave rationally.

Table III. Mean total WTP (in NOK) for the sub-samples receiving one or two valuation questions, and test statistics for the hypotesis (H-1).

Variable	Mean	t-value
Total WTP for sub-sample with one valuation question	1144	
Total WTP for sub-sample with ιwo valuation questions	1139	$T_{H\text{-}1} = 0.0041$

4.3. Econometric Analysis of the Data

Internal consistency is also tested using regression analysis. The dependent variable in the analysis is total WTP, estimated either directly or in a sequence. Three main categories of independent variables are included: (i) variables describing the respondents' current health condition, (ii) attitude-variables, and (iii) socio-economic and demographic variables. We also included dummy variables for the two different splits. The dummy for receiving two WTP questions is used to test the hypothesis (H-1).

The WTP variable does not take values below zero, and has a positive density at a zero WTP. In this case, Ordinary Least Squares (OLS) will yeald biased estimates (Maddala 1988). We have therefore applied both a TOBIT model and a Cragg specification. The main idea in the TOBIT model is that the WTP is a continuous variable, but we only observe those cases with a positive WTP. For all observations where the real WTP (y^*) is negative or equal to zero, the observed WTP (y) is zero. For all observations where the WTP is positive, it is assumed that we can observe the true WTP. Taking account of the fact that the density at a zero WTP is positive, the expression for the expected WTP in a TOBIT model is given by:

$$E(y_i) = \Phi\left(\frac{\beta' x_i}{\sigma}\right)(\beta' x_i + \sigma\lambda_i),$$

where

$$\lambda_i = \frac{\phi(\beta' x_i/\sigma)}{\Phi(\beta' x_i/\sigma)} \quad (3)$$

ϕ () is the standard normal density function, and Φ () is the standard normal probability. See Greene (1993) or Maddala (1983) for a more detailed description of the TOBIT model.

The TOBIT model is a special case of a specification suggested by Gragg (1971). Here, the structure of the decision of whether the respondents have a positive WTP or not is assumed to be independent of the structure of the WTP for those respondents who are willing to pay. The Cragg model consists of two independent regressions: (a) one probit model for stating a zero WTP:

Table IV. Estimation results from: (1) the TOBIT model and (2) the Cragg specificaton: (*a*) a probit model for the probability of stating a positive WTP, and (*b*) a truncated regression model for respondents with a positive WTP. The probabilities of falsely rejecting the zero hypothesis of no effect (significant level) are given in parentheses

Model Variable	(1) TOBIT model	(2) Cragg specification (*a*) Probit model	 (*b*) Truncated model
Constant	−4503 (0.000)	−1.113 (0.000)	−24631 (0.022)
Household income (10.000 NOK)	44 (0.000)	0.009 (0.002)	214 (0.005)
Age	−33 (0.000)	−0.009 (0.000)	−128 (0.062)
Major city	915 (0.003)	0.315 (0.001)	760 (0.672)
University degree	1097 (0.001)	0.235 (0.022)	4781 (0.032)
Concerned with the environment	1525 (0.002)	0.430 (0.003)	4198 (0.349)
Government doing too much	−4996 (0.007)	−1.360 (0.008)	−11932 (0.616)
Governing doing too little	1428 (0.000)	0.377 (0.000)	4360 (0.073)
Positive to electric cars	1767 (0.002)	0.513 (0.003)	6398 (0.186)
Negative to electric cars	−1964 (0.115)	−0.625 (0.079)	9753 (0.406)
Minor health problems	1142 (0.000)	0.328 (0.000)	2327 (0.254)
Asthma, allergy, bronch., emfys.	745 (0.006)	0.220 (0.008)	1981 (0.266)
Electric car scenario	−1410 (0.014)	−0.348 (0.037)	−7724 (0.122)
Two valuation questions	**107 (0.687)**	**0.002 (0.977)**	**1421 (0.340)**
Variation in the WTP (σ^2)	3694 (0.000)		6589 (0.000)
Log-likelihood	−4794	−654	−4109

$$\begin{aligned} \text{Prob}(y_{i^*} > 0) &= \Phi(\gamma' x_i) \qquad y_i = 1 \ \text{ if } \ y_{i^*} > 0 \\ \text{Prob}(y_{i^*} \leq 0) &= 1 - \Phi(\gamma' x_i) \quad y_i = 0 \ \text{ if } \ y_{i^*} \leq 0 \end{aligned} \tag{4}$$

(b) one truncated regression model for respondents stating a positive WTP:

$$E(y_{i^*} | y_i = 1) = \beta' x_i + \sigma \lambda_i. \tag{5}$$

The TOBIT model is a special case of the Cragg model when $\gamma = \beta/\sigma$. See Cragg (1971) or Greene (1993) for more information.

The estimates from these regressions are presented in Table IV: the TOBIT estimates in the first column, and the results from the Cragg specification in the second and the third column. In this paper, we will only comment on the results relevant for the testing of our hypothesis (H-1). The testing of (H-1) is equal to a test of whether the coefficient for receiving two valuation questions in significantly different from zero. Looking at the results from the TOBIT model, we find that respondents who had to answer the WTP questions in a sequence state a higher total WTP than those respondents who were only asked one WTP question. However, this difference is far from significant with a probability of rejection H_0 when H_0 is true at almost 69%. Looking at the results from the Cragg specification, the same

Table V. Estimation results from: (1) the TOBIT model and (2) the TOBIT model corrected for heteroscedasticity: (*a*) the effect on the expected WTP, and (*b*) the effect on the variation in WTP. The probabilities of falsely rejecting the zero hypothesis of no effect (significant level) are given in parentheses.

Model / Variable	(1) Uncorrected TOBIT model	(2) Corrected TOBIT model (*a*) E(WTP)	(*b*) $\sigma_i^2 = \text{var}_i$
Constant	−4503 (0.000)	−346 (0.612)	
Household income (10.000 NOK)	44 (0.000)	8 (0.453)	0.001 (0.000)
Age	−33 (0.000)	−43 (0.000)	−0.002 (0.276)
Major city	915 (0.003)	370 (0.270)	−0.018 (0.799)
University degree	1097 (0.001)	617 (0.095)	0.146 (0.053)
Concerned with the environment	1525 (0.002)	357 (0.449)	0.165 (0.182)
Government doing too much	−4996 (0.007)	−528 (0.697)	−1.315 (0.198)
Government doing too little	1428 (0.000)	616 (0.052)	0.100 (0.227)
Positive to electric cars	1767 (0.002)	494 (0.381)	0.233 (0.205)
Negative to electric cars	−1964 (0.115)	−378 (0.775)	0.109 (0.760)
Minor health problems	1142 (0.000)	435 (0.152)	0.157 (0.046)
Asthma, allergy, bronchitis, emphyfemta	745 (0.006)	211 (0.476)	0.068 (0.325)
Electric car scenario	−1410 (0.014)	−367 (0.511)	−0.306 (0.093)
Two valuation questions	**107 (0.687)**	**68 (0.813)**	**0.039 (0.535)**
Variation in the WTP (σ^2)	3694 (0.000)		2058 (0.000)
Log-likelihood	−4794		−4787

result is derived. Receiving two valuation questions did not affect the probability of stating a positive WTP nor the expected WTP for the respondents stating a positive WTP. The probabilities of falsely rejecting the hypothesis when it is true are 98% and 34% respectively. These tests strongly support the hypothesis (H-1).

However, the maximum likelihood estimates of the TOBIT model may be inconsistent if heteroscedasticity occurs (Maddala and Nelson 1975; Hurd 1979; Brown and Moffit 1982). Peterson and Waldman (1981) suggest a way to estimate a TOBIT model corrected for heteroscedasticity. The main idea is to replace σ in equation (3) with σ_i, where $\sigma_i^2 = \sigma^2 \exp(\alpha\, x_i)$, and estimate both β and α simultaneously (Greene 1993).

The results from this estimation are presented in Table V. The first column shows the TOBIT estimates without correction for heteroscedasticity. The second and third column presents the result from the corrected model. Receiving two valuation questions had a positive effect both on the corrected expected WTP and the variation in the WTP answers, but neither of these effects were significant. The probability of falsely rejecting the hypothesis is 81% and 54% respectively. Thus, we do not reject the hypothesis (H-1). We therefore conclude that the respondents are able to answer the valuation questions consistently. This was the same impression derived from analysing the descriptive statistics.[8]

5. Concluding Remarks

The results from our survey show that a sequential valuation procedure may create considerable and significant ordering effects and/or part-whole biases. The main reason for such biases seems to be that the respondents are given imperfect information about the valuation problem during the valuation sequence. We would like to stress the importance of perfect information on the validity of the results from a CVM survey. First, if the aim of the survey is to value a particular effect of one project (or a particular project from a policy agenda), without informing the respondents about all effects of the project (all projects in the agenda), it will create embedding effects for the reasons discussed in Randall and Hoehn (1989), and potential part-whole biases. Second, if the aim is to value all effects of a project (all projects in the agenda) in a valuation sequence, and additional information is given to the respondents during the sequence, it may create considerable unwanted ordering effects and potential part-whole biases. These imperfections in information can even lead to internal inconsistency, as in Samples and Hollyer (1990). A way to solve this information problem in the sequential valuation setting may be to allow the respondents to reconsider their initially stated WTP after all information has been revealed. Alternatively, one can explain everything to the respondents up front: all the goods that are going to be valued, whether all the goods are effects of one project or not, and that they are going to be asked to value these goods in a sequence. This may be too much information at once, and the task of valuation may in such contexts be quite demanding for some respondents. Thus, a one-short, holistic valuation of all goods may be preferable to a sequential valuation. If a decomposition is required, the top-down approach can be applied. This approach may increase the variation in the estimates, but it will not yield biased estimates as the sequential decomposition method (the bottom-up approach) will do if the real decision problems is of a simultaneous nature.

Acknowledgments

The author is grateful to Gunnar Rongen, Jon Strand, and the referees of this journal for valuable comments on previous versions of the paper. This paper is a part of a project called "Valuation and Use of Environmental Costs" under the SAMMEN-programme, financed by the Research Council of Norway.

Appendix: Likelihood Ratio Tests

In this Appendix, we will test the TOBIT model against two alternate specifications: the Cragg specification and a TOBIT model corrected for heteroscedasticity. This is done by applying a Likelihood ratio test. For more information about Likelihood ratio tests, see Greene (1993).

The Cragg specification

A TOBIT model can be seen as a restricted model of the Cragg specification under the hypothesis; H_0: $\gamma = \beta/\sigma$. Thus, the chi-square distributed test statistic for the likelihood ratio test is given by:

$$LR = -[\ln L_T - (\ln L_P + \ln L_{TR})] = 62 \sim \chi^2_{(13d.f.)}, \quad \chi^2_{chr.(0.95;13)} = 22.362.$$

Since the observed test statistic exceeds the critical value for a 5% test, we reject the hypothesis H_0 and thus prefer the Cragg specification. For more information about this test, see Green (1993), 701.

A TOBIT-model corrected for heteroscedasticity

A TOBIT model can be seen as a restricted model of a TOBIT model corrected for heteroscedasticity, under the hypothesis of homoscedasticity; H_0: $\alpha = 0$. Thus, the chi-square distributed test statistic for the likelihood ratio test is given by:

$$LR = -2[\ln L_T - \ln L_C] = 14 \sim \chi^2_{(13d.f.)}, \quad \chi^2_{chr.(0.95;13)} = 22,362.$$

Since the observed test statistic does not exceed the critical value for a 5% test, we fail to reject the hypothesis H_0 and thus prefer the TOBIT model. For more information about this test, see Green (1993), p. 699.

Notes

[1] Hoehn and Randall (1989) discuss problems that occur if a project is valued on its own when it is part of a policy agenda, but the same conclusions can be drawn if a project has more than one effect on the populations welfare.

[2] Randall (1991) denotes this 'the independent valuation and summation' (IVS), and Hoevenagel (1994) 'the added total value'.

[3] Randall (1991) denotes this 'the total value' (TV), and Hoevenagel (1994) 'the direct total value'.

[4] Here we only consider the adding-up test adjusted for income effects (see Diamond and Hausman 1991: 52, footnote 11).

[5] The utility function is assumed to satisfy the general assumptions. It is also assumed to be declining and concave in K, and linear in π.

[6] This is true as long as K and π are substitutes. If they are complements, the substitution effect is assumed to exceed the income effect.

[7] We define Bounded Rational as not capable of maximizing utility for a given income and the given set of information at the decision point.

[8] See the Appendix for the results from likelihood ratio tests between the different models.

References

Brown, R. and R. Moffitt (1982), 'The Effect of Ignoring Heteroscedasticity on Estimates of the Tobit Model', *Mimeo, University of Maryland, Department of Economics, June 1982.*

Clench-Aas, J., S. Larssen, A. Bartonova and M. Johnsrud (1989), 'Virkninger av luftforurensninger fra veitrafikk på menneskers helse. Resultater fra en undersøkelse i Vålerenga/Gamlebyen-området i Oslo, 1987', *NILU OR-70/89*, Ref. O-8638. Norway.

Diamond, P. A. and J. A. Hausman (1994), 'Contingent Valuation: Is Some Number Better than No Number?', *Journal of Economic Persectives* **8**, 45–64.
Goldsmith, J. R. and L. T. Friberg (1977), 'Effects of Air Pollution on Human Health I: Stern', *Air Pollution*, 3rd edn., New York: Academy Press, pp. 457–610.
Greene, W. H. (1993), *Econometric Analysis*, Macmillan, New York.
Hoehn, J. P. and A. Randall (1989), 'Too Many Proposals Pass the Benefit Cost Test', *The American Economic Review* **79**, 544–551.
Hoevenagel, R. (1994), *The Contingent Valuation Method: Scope and Validity*, Vrije University, Amsterdam.
Hurd, M. (1979), 'Estimation of Truncated Samples When There Is Heteroscedasticity', *Journal of Econometrics* **11**, 247–58.
Kahneman, D. and J. I. Knetsch (1992), 'Valuing Public Goods: The Purchase of Moral Satisfaction', *Journal of Environmental Economics and Management* **22**, 57–70.
Maddala G. S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.
Maddala, G. S. (1988), *Introduction of Econometrics*, Macmillan, New York.
Maddala, G. S. and F. Nelson (1975), 'Specification Errors in Limited Dependent Variable Models', *Working Paper 96*, National Bureau of Economic Research, Cambridge.
Perry, G. B., H. Chai, D. W. Díckey, R. H. Jones, R. A. Kinsman, C. G. Morill, S. L. Spectot and P. C. Weiser (1982), 'Effects of Particulate Air Pollution on Asthmatics', *American Journal of Public Health* **73**, 50–56.
Peterson, D. and D. Waldman (1981), 'The Treatment of Heteroscedasticity in the Limited Dependent Variable Model', *Mimeo, University of North Carolina, Chapel Hill, November 1981*.
Randall, A. (1991), 'Total and Nonuse Values', *Measuring the Demand for Environmental Quality*, in J. B. Barden and C. D. Kolstad, eds., North-Holland Amsterdam.
Samples, K. C. and J. R. Hollyer (1990), 'Contingent Valuation of Vildlife Resources in the Presence of Substitutes and Complements', in R. L. Johnson and G. V. Johnson, eds., *Economic Valuation of Natural Resources*, Westview Press.