TI 2008-053/3
Tinbergen Institute Discussion Paper

The Information Content of a Stated Choice Experiment – A New Method and its Application to the Value of a Statistical Life

Jan Rouwendal[1]
Arianne de Blaeij[2]
Piet Rietveld[1]
Erik Verhoef[1]

[1] *VU University Amsterdam, and Tinbergen Institute,*
[2] *LEI, The Hague.*

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

The information content of a stated choice experiment

A new method and its application to the value of a statistical life

Jan Rouwendal*, Arianne de Blaeij, Piet Rietveld* and Erik Verhoef***

* Department of Spatial Economics, VU University, Amsterdam
** LEI, Den Haag

This version: May 16, 2008

Keywords: stated preferences, value of a statistical life
JEL Classification Codes: C81, D12, D61, R41

Abstract
This paper presents a method to assess the distribution of values of time, and values of statistical life, over participants to a stated choice experiment, that does not require the researcher to make an a priori assumption on the type of distribution, as is required for example for mixed logit models. The method requires a few assumptions to hold true, namely that the valuations to be determined are constant for each individual, and that respondents make choices according to their preferences. These assumptions allow the derivation of lower and upper bounds on the (cumulative) distribution of the values of interest over respondents, by deriving for each choice set the value(s) for which the respondent would be indifferent between the alternatives offered, and next deriving from the choice actually made the respondent's implied minimum or maximum value(s). We also provide an extension of the method that incorporates the possibility that errors are made. The method is illustrated using data from an experiment investigating the value of time and the value of statistical life. We discuss the possibility to improve the information content of stated choice experiments by optimizing the attribute levels shown to respondents, which is especially relevant because it would help in selecting the appropriate distribution for mixed logit estimates for the same data.

1 Introduction

Discrete choice analysis, of both stated and revealed preference data, is an important tool of transportation research. The mixed logit model has recently become an important tool for this type of analysis (see e.g. Hensher et al., 2005). A main advantage of this model, over more conventional alternatives such as the multinomial and nested logit models, is that it can deal with variations in preferences over respondents by allowing estimated parameters to follow certain distributions. The multinomial logit model can of course explicitly incorporate taste variation by relating it to observed characteristics of the respondent, but there is often substantial remaining heterogeneity within classes defined by observed characteristics. It is therefore not too surprising that treating the taste parameters as random variables, as in the mixed logit model, is important in many cases. Moreover, the mixed logit model does not suffer from some other limitations that are inherent to the specification of standard and nested logit model (see McFadden and Train, 2000).

The specification of a mixed logit model requires the choice of a particular type of distribution function for the random parameters. Theory usually offers little guidance for this choice, which is therefore often guided by convenience, *a priori* plausibility, or even by something as pragmatic as the convergence of model estimations. Because central estimates of parameters of interest often vary over specifications, this is somewhat problematic. Because alternative model formulations are often non-nested, the selection of the best model is not straightforward. One could apply a flexible formulation that is able to approximate any arbitrary distribution of the random coefficients. This is done in the latent class approach, which is popular especially in marketing (see Kamakura and Rusell, 1989). However, in many applications a mass point distribution is not intuitive, and the choice of the appropriate number of groups is often somewhat arbitrary (see Wedel et al. 1999 for a discussion of these and related issues).

It therefore seems desirable to have a method that would enable a researcher to investigate the distribution of parameters of interest, like the value of a statistical life (*vosl*) or the value of time (*vot*), without having to make *a priori* assumptions about the functional form of the distribution of the random parameters. This paper proposes a method for exploring the distribution over individuals of the *vosl* and *vot*, or similar variables, under some minimal *a priori* assumptions. These assumptions are, first, that these marginal valuations are individual-specific constants, at least over the ranges considered; and second, that the choices made by the respondents reveal their true preferences.

These two assumptions allow one to consider each response to a dichotomous choice situation as a revelation of a lower or upper bound on the valuation of interest. For example, if a respondent prefers a trip that takes 10 minutes longer but costs 1 Euro less over an alternative trip that is, besides price and travel time otherwise identical, one could conclude that this respondent's value of time is not above 6 Euros per hour. If she chooses the alternative, it is not below 6 Euro's per hour. If the alternatives are defined by more than two attributes, as in our empirical case, every observed choice still produces an inequality characterizing the individual's preferences, and therefore defines a bound on the feasible set of combinations of marginal valuations that are consistent with the individual's choices (maintaining the assumptions that one of the attributes is monetary, and that the marginal valuations are constants). For the data analyzed here, the half-spaces can be pictured as part of a two-dimensional diagram with the *vot* and the *vosl* on its axes. Geometrically, every choice situation thus divides the space of relevant marginal valuations into two "half-spaces", and by making a choice the respondent reveals to which of these two half-spaces his marginal valuations belong (which is why we will refer to this method as the 'half-space method'). A sequence of choices will then, with each successive choice, typically further narrow down the possible range in which the marginal valuation(s) can lie, and will thus eventually define a lower and an upper bound for every valuation. Provided the individual's choices are mutually consistent (under the assumption of constant marginal valuations), the former is below the latter. Combining these bounds across individuals, one can obtain aggregate distributions for the lower and upper bounds for the valuation of interest, and for example compare these with the distributions obtained for various specifications of mixed logit models.

These bounds are, of course, more informative of the distribution of point estimates of marginal valuations when these bounds are closer. The closeness of bounds will be shown to depend on the statistical design of the stated choice experiment. The half-space method can therefore also be useful in the design of stated choice experiments, by suggesting how to focus the choice experiment on the relevant ranges of the parameter(s) of interest.

We discuss the method using data that were collected with the prime objective to investigate the value of statistical life (*vosl*) in road traffic for Dutch citizens,[1] producing value of time (*vot*) estimates as an intended by-product. The *vot* and *vosl* distributions derived show a

[1] See Ashenfelter (2006) for a recent literature on the *vosl*.

substantial amount of variation. The upper and lower bounds that follow from the half-space method are informative: some points of this distribution function are indicated exactly, and for a range of values the upper and lower bounds are close to each other. Earlier analyses of these same data contributed to the determination of a *vosl* that is currently used in Dutch traffic safety policy (see Wesemann *et al*., 2005). Mixed logit models were estimated with normal and lognormal distributions for the taste parameters of interest, namely the toll to be paid and the number of fatalities per million trips. The normal distributions have the disadvantages of postulating that part of the population have a negative *vosl*; and, because the *vosl* is the ratio between normal distributed coefficients, that its distributions can be peculiar (see Meijer and Rouwendal, 2006). Lognormal distributions for both parameters have a relatively 'fat tail', which results inevitably in a fat tail of the estimated distribution of the *vosl*. This raises the concern that this is an artifact of the chosen specification.

These problems are illustrative for applications of the mixed logit model and underline the need for a method to investigate the distribution of the parameters of interest without making strong *a priori* assumptions. To elaborate the point, Sonnier, Ainslee and Otter (2003) and Train and Weeks (2004) have recently compared mixed logit specifications estimated in preference space (where the willingness to pay is computed as the ratio of two random parameters) with specifications estimated in willingness to pay space (where the distribution of the willingness to pay is specified directly). Their tentative conclusion is that "... models in preference space fit the […] data better than […] models in wtp space, but provide unreasonably large variances in wtp" (Train and Weeks, 2004). For the data we use here, we estimate mixed logit models in both spaces, and we obtain results that are qualitatively similar to those of these papers. The distributions of the *vosl* implied by the two specifications are both between the upper and lower bounds from the half-space method, so that the method cannot be used to choose between them.

The half-space method also provides a check for the mutual consistency of the answer choices made by the respondents: if an individual's lower bound is above the upper bound, the choices can be characterized as inconsistent (under the maintained assumption of constant marginal valuations). Probably such inconsistencies are caused (partly or completely) by erroneous choices, which suggests that the chances of providing a completely consistent sequence of choices decreases with the number of choices that has to be made. In our data, 10 choices had to be made by all respondents, and we find that approximately 35% of the choice sequences are

inconsistent. Introduction of a simple error generating mechanism into the model allows us to also incorporate these inconsistent choices into the analysis. The estimated probability that a respondent's choices are in accordance with his or her preferences lies between 90 and 95%. The upper and lower bounds for the distribution function of the *vosl* implied by this model are close to those computed for the consistent respondents only, but they are in general somewhat higher.

The paper is organized as follows. The next section provides a brief discussion of the data we use. Section 3 introduces the non-parametric half-space method for investigating the distribution of the parameters of interest. Section 4 investigates the implied upper and lower bounds for the distribution of the *vosl*. Section 5 deals with the incorporation of inconsistent choices. In section 6 we compare the implied distributions of mixed logit estimates in preference space and willingness to pay space with the bounds derived earlier. Section 7 briefly discusses the results of a similar analysis with respect to the *vot* on the same data. Section 8 concludes.

2 The data

The data we will use were gathered as part of a larger internet survey carried out by a specialized Dutch bureau (Intomart). The information used here refers to a number of stated-choice questions that were formulated in order to investigate the respondent's valuation of changes in traffic fatality probability. Each respondent was asked to imagine that (s)he has to make a trip from A to B by car, while there are no other persons in the car. Two roads can be used for the trip, which may differ in three attributes: toll, the probability of a fatal accident, and travel time. It was stressed that the two roads differ only in these three attributes. The main interest of the survey was to investigate the marginal valuation of traffic fatality probabilities (expressed as the *vosl*), and travel time was included primarily to facilitate a comparison of the results with earlier travel time valuation studies. This was considered desirable since no prior *vosl* studies had been carried out earlier in the Netherlands, and the plausibility of estimates from this study might partly be based on whether the *vot* outcomes are within reasonable ranges (as they turned out to be).

A simple orthogonal main-effects only design was used for this study, where each respondent made 9 choices (one choice was repeated as a 10^{th} choice, to check consistency), and the full design was split into 5 blocks.[2] Attribute levels within each block were generic over respondents: there was pivoting of a respondent's personal trip. Strictly dominant choices were

[2] Some attribute levels were changed in order to avoid dominated alternatives, see e.g. Rizzi and Ortuzar (2003).

not included in the design: the differences in the toll and the number of fatalities were always of the opposite sign. The travel time attribute did not always differ between the two roads, but if it did the difference always had the same sign as that between the number of fatalities. This means that the choice situations posed to the respondents always implied that they had to pay for additional safety, possibly in combination with travel time savings.

The three attributes were specified as follows. The toll is the price per trip in Dutch guilders (Dfl),[3] and varies between Dfl 2.50 and Dfl 12.50. The travel time varies between 50 minutes and 1 hour. Road safety is indicated by the annual number of fatalities on the road, which varies between 12 and 36. The respondents were informed that the total annual number of trips made on the road is 18 million, which means that the lowest number of fatalities (12) corresponds to the average safety level on Dutch roads. The *vosl* figures provided in this paper are the ratio of the marginal utility of this objective "generic" accident probability and the marginal utility of money (toll). It may be that respondents believe they are at lower or higher risk than the average road user. If so (this was not investigated), the correct interpretation of our *vosl* figures involves the willingness to pay to improve average, not individual safety. The English translation of the exact stated choice question is provided in the Appendix A.

Table 1 Basic information about the data

Group	Number of respondents	Different choices in 2 and 10
1	207	33
2	220	36
3	211	20
4	215	41
5	202	29
Total	1055	159

Note. The number of choice situations for which there is no difference in travel time between the two alternatives for groups 1-5 is 3,3,1,3 and 3, respectively.

There were 1055 respondents. There are no missing data since respondents had to provide an answer in order to be able to proceed to the end of the questionnaire, and consequently to receive their payment. The necessity to give a response may have had a deteriorating effect on the

[3] 1 Dutch guilder was .45 euro.

quality of the responses and this makes it desirable to have a reliability check. For this reason, the second and the tenth choice situations were made identical for all respondents.

The numbers of respondents in each of the five blocks, referred to as 'groups' (of respondents) in the sequel, are given in Table 1. The table also provides information on the number of respondents who made different choices in situations 2 and 10. Approximately 15% of the respondents did so.[4]

3 The half-space method

The basic hypothesis behind stated choice analysis is that the choices made by respondents reflect their preference ordering over all alternatives. These preferences are usually described by means of a utility function u, which has the attributes x of the alternatives as its arguments, and possibly also the respondent's (observed and unobserved) characteristics z.

$$u = u(x; z) \qquad (1)$$

The preferences of the respondents with respect to the roads between which they have to choose in our experiment depend on three road characteristics: the toll to be paid, the travel time, and the fatality probability on the road. The marginal rate of substitution between travel time and toll is the value of time (*vot*), and that between the fatality probability and toll is the value of statistical life (*vosl*). The former is expressed in money per unit of time, and the latter in money per 'unit of probability'. Because a 'unit of probability' corresponds to the extreme difference between a completely certain non-occurrence and a completely certain occurrence of an incident, it is important to emphasize that the *vosl* is the marginal willingness to pay, referring to marginal changes in fatality probabilities. What is valued are infinitesimally small changes in fatality probabilities; it is the units in which these are measured that may make to willingness to pay *seem* to refer to the avoidance of a certain death – which it certainly does not, because the difference between probabilities of 0 and 1 is of course definitely not 'marginal'.[5]

[4] Loomes et al. (2002, p. 103) indicate that an order of 20 to 30 per cent for inconsistencies of this kind is typical in the literature.

[5] Instead, the *vosl* aims to reflects the individual's marginal valuation of (infinitesimally small) changes in risk – it is the rather arbitrary choice of probability "units" (with units chosen such that the probability may vary between 0 and 1) that express this valuation in a unit that represents a certain fatality. Just as there is no reason to impose an income constraint on an individual's marginal willingness to pay for apples when the units of measurement changes from single apples to the yearly global apple production, there is no reason to impose an income constraint on the *vosl* when it is used to evaluate the benefits of small changes in fatality probabilities.

The marginal rates of substitution, in turn, are the ratios between the marginal utilities of time and toll for *vot*, or of fatality probability and toll for *vosl*. The simplest illustration is for an indirect utility function that is linear in the three attributes:

$$u = \beta_{toll} \cdot \tau + \beta_{prob} \cdot P + \beta_{time} \cdot T \qquad (2)$$

where τ is toll, P is fatality probability, and T is travel time. The *vosl* and *vot* implied by this utility function are:

$$vosl = \frac{\beta_{prob}}{\beta_{toll}}; \qquad vot = \frac{\beta_{time}}{\beta_{toll}}. \qquad (3)$$

Conventional discrete choice models usually estimate parameters β for a utility function resembling (2) but with a random term added, and next determine the *vot* or *vosl* according to (3). Depending on the model formulation chosen, the estimated parameters β in (2), and the implied marginal valuations *vot* or *vosl* in (3), may or may not be allowed to vary over respondents in accordance with observed or unobserved heterogeneity.

For the half-space method explored in this paper, we immediately allow for the *vot* and *vosl* to vary over individuals, but do assume that they are constant across choices for each respondent individually. This allows us to pool the observations for each individual to determine a lower and upper bound on that individual's marginal valuations *vot* and *vosl*. In reality, the individual's *vot* and *vosl* of course need not be constant across choices, in particular if the marginal valuations (β) vary with attribute levels. Keeping the marginal valuations constant has the great advantages of producing a single *vot* and a single *vosl* for each individual, so that we can determine straightforward unidimensional distributions of *vot* and *vosl* over individuals. A main disadvantage is that we introduce a second source of seemingly inconsistent responses, besides the sheer possibility of ticking the wrong button on a computer screen, possibly because of reduced attention or hasty reading. A combination of choices that would be inconsistent with a constant *vot* and *vosl* for an individual, may be perfectly consistent with valuations that vary with attribute levels. Nevertheless, a linear approach will then still give a lower and upper bound that is representative for the individual for the particular range of attribute levels covered by the choice sets.

Now consider a choice set with two alternatives *i* and *j* presented to an individual, and denote the differences between attribute levels, for alternative *i* minus alternative *j*, with Δ. With

constant marginal valuations *vot* and *vosl*, the individual will be indifferent between the two alternatives if:

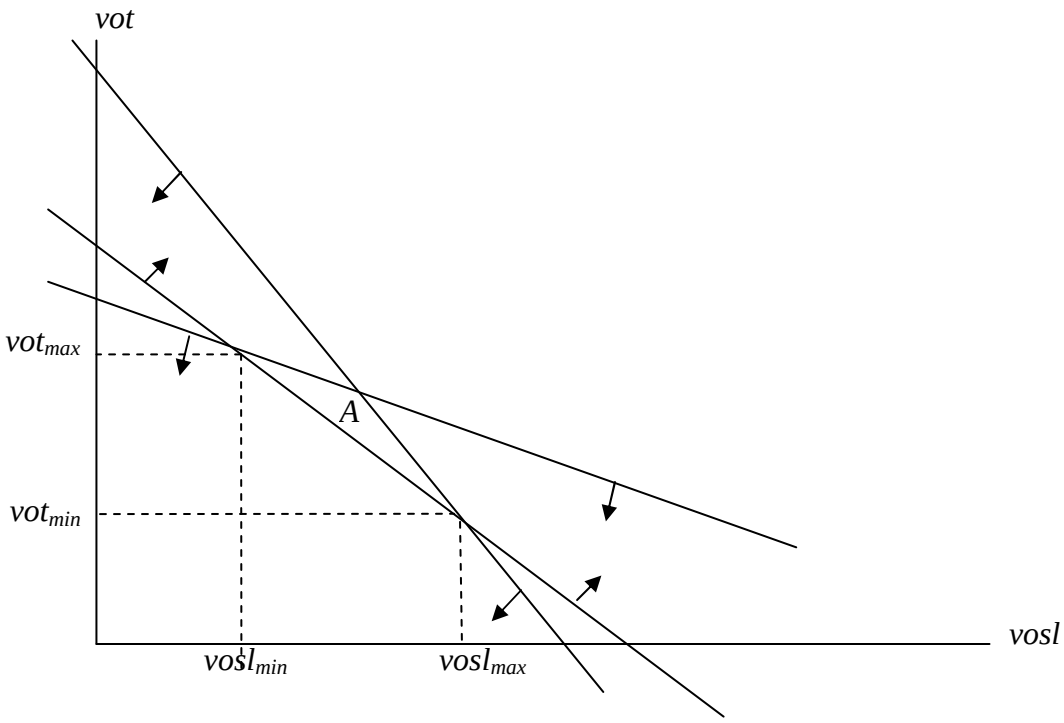
$$\Delta\tau + vosl \cdot \Delta P + vot \cdot \Delta T = 0 \quad \Leftrightarrow \quad vosl = -\frac{\Delta\tau}{\Delta P} - vot \cdot \frac{\Delta T}{\Delta P} \tag{4}$$

Equation (4) defines an affine equation in the *vot*-*vosl* space, which connects values of *vot* and *vosl* for which the individual would be indifferent between the two alternatives. The actual choice made by the individual therefore reveals on which side of this line her combination of *vot* and *vosl* is to be found. A numerical example may help: consider an individual who prefers a trip that has a toll exceeding the other toll by 10, while the travel time is 1 hour less and the accident risk is 1:1 million smaller. The individual may then for example have a *vot* of 0 and a *vosl* of at least 10 million. She may also have a *vosl* of 0 and a *vot* of at least 10. There are of course countless possible combinations of *vot* and *vosl* that would be consistent with the observed choice. But, we do know that the *vot* is at least 10 – *vosl* / 1 million. Her true combination of *vot* and *vosl* is therefore not to the south-west of this line when plotted in *vosl*-*vot* space – and this is the 'line of indifference' defined by equation (4) for this particular numerical example. Of course, if the individual chooses the cheaper but slower and less safe alternative, we know that her combination of *vot* and *vosl* is not to the north-east of this same line. We therefore know, from the observed choice, in which of the half-spaces defined by the line of indifference (4) the respondent's *vosl*-*vot* combination lies.

When the respondent makes a number of choices, each of them defines such a half-space. The respondent's combination of *vot* and *vosl* must then lie somewhere in the intersection of all these half-spaces; at least if, as we assume, *vosl* and *vot* are individual specific constants. These intersections, in turn, imply lower and upper bounds on the respondents' *vot* and *vosl*, which in turn can be combined across respondents to find the simultaneous distribution of these variables.

Figure 1 provides a graphical example for an individual who has made three choices, with the implications for the *vosl*-*vot* combination represented by the arrows attached to each 'line of indifference'. The three choices together imply that the true combination of *vot* and *vosl* must be in triangle *A*, the intersection of the three half-spaces, so that for this example we can identify finite minima and maxima for both *vot* and *vosl*.

Figure 1. The half-space method illustrated with three 'lines of indifference'



It will be clear that this procedure works only when the respondent's choices are mutually consistent. It is not hard to imagine a fourth line in Figure 1, to the north-east of triangle *A* and with arrows pointing to the north-east, that would leave the intersection of half-spaces implied by this individual's choices is empty. One possibility for such an inconsistency to arise in our dataset is when different answers are provided to the identical questions 2 and 10 – although this could still be taken as a sign that the respondent's true combination of *vot* and *vosl* happens to be exactly *on* the line of indifference implied by the choice set. Inconsistencies may of course also arise for respondents who do provide the same answer to questions 2 and 10, and Table 2 shows that the total number of respondents with inconsistent choices is approximately twice as large as the number that made different choices for these test questions. Approximately 30% of the respondents made choices that are mutually inconsistent in the sense described above (the intersection of half-spaces is empty) if we maintain that *vot* and *vosl* are constant for each individual. In the remainder of this section and in the next one we ignore these respondents, but we will return to them in section 5.

Table 2 Consistent choices

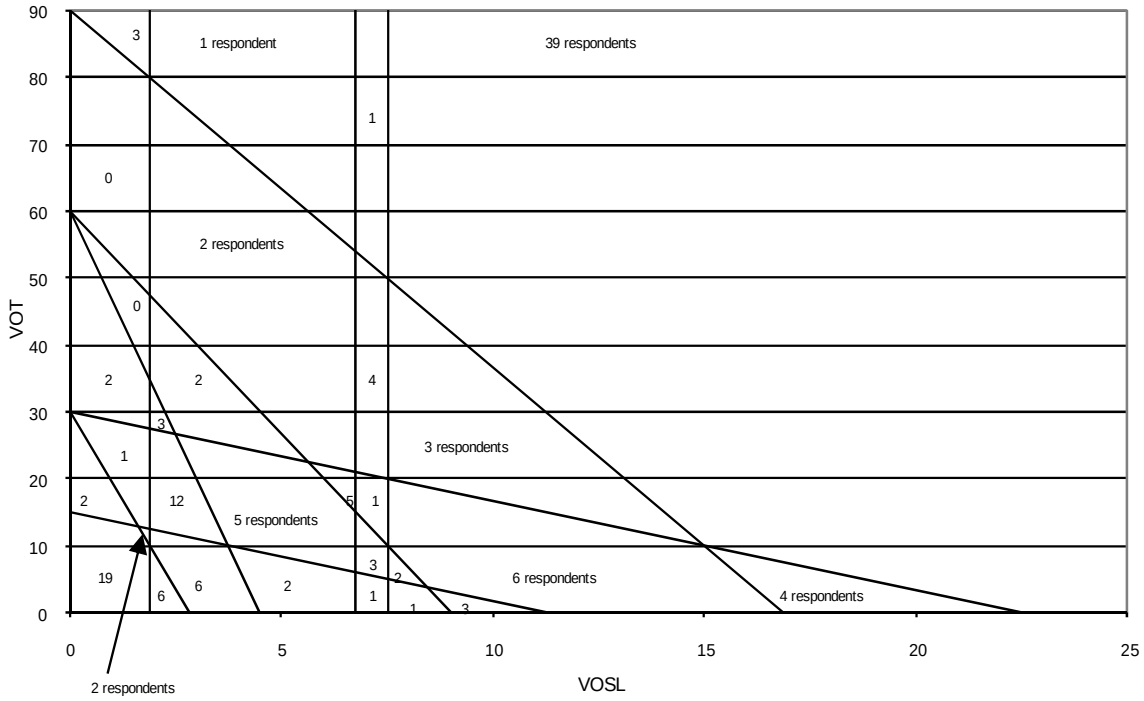
Group	*n*	Consistent	Inconsistent
1	207	117	90
2	220	140	80
3	211	118	93
4	215	136	89
5	202	136	76
Total	1055	664	391

Figure 2 shows the lines of indifference, and the implied possible intersections of half-spaces in which a respondent with consistent choices can end up, for one of the five groups in our experiment. We only show results for group 4; comparable diagrams for the other groups are available upon request. The diagram shows that some of these intersections are relatively small, defining relatively small intervals for *vot* or *vosl*, while in other cases these intervals are wide and, unavoidably, in a number of cases unbounded.

The vertical lines of indifference refer to choices between alternatives with equal travel times. Such choices imply a unique critical value of the *vosl*, and are therefore especially informative for the present purposes. We can for example see that among the 136 respondents present in Figure 2, 27 (19+2+2+1+3) have a *vosl* below Dfl 1.875 mln – the value implied by the first vertical line of indifference. The critical value of the *vosl* defined by other choices depends on the *vot*, which diminishes their informational contents for the *vosl* but of course raises it for the *vot*.

Figure 2 thus summarizes all information that the choice experiment provides about the values of time and the values of a statistical life for respondents in group 4, under the two – arguably minimal – assumptions that choices reveal true preferences and that *vot* and *vosl* are individual-specific constants.

Figure 2. Lines of indifference and occupation of half-spaces for consistent sets of choices, group 4 (vot in Dfl, vosl in mln. Dfl)



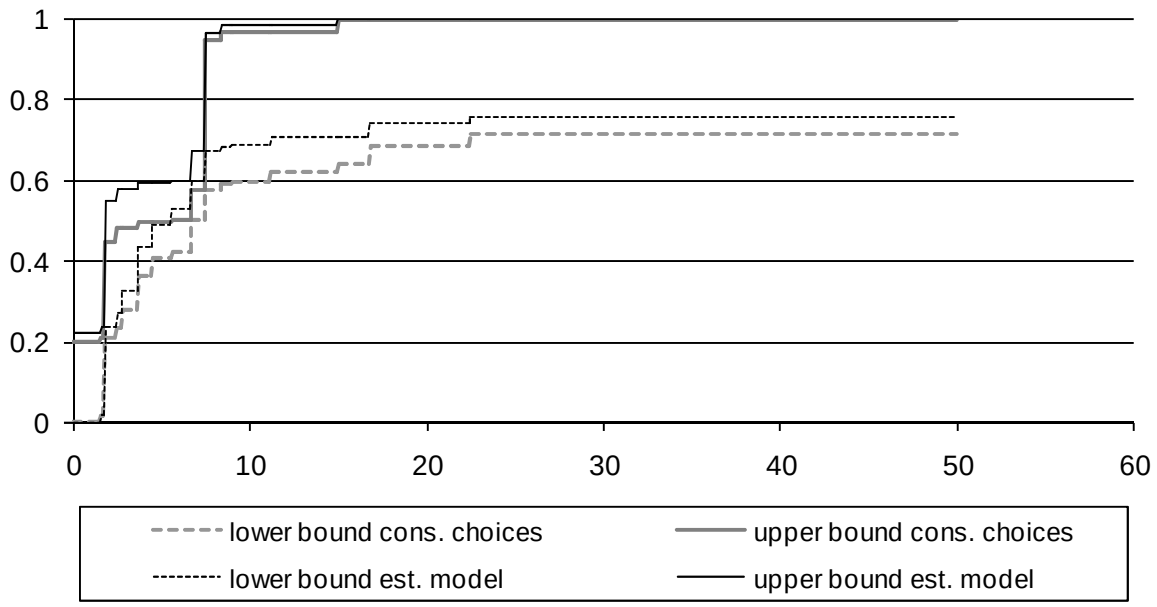
4 Upper and lower bounds on the VOSL

Figure 2 strongly suggests that the values of time and of a statistical life differ over the respondents. There is not a single value of these variables that is (approximately) applicable to everyone, but instead there seems to be a range of values.

Since every consistent respondent can be allocated into a particular area defined by the lines of indifference in Figure 2, the interval to which an individual's *vosl* belongs is defined by the lowest and highest values of the *vosl* (the latter possibly being infinite) that belong to that area; compare also Figure 1 (the same holds for a minimum and maximum *vot*). These upper and lower bounds of the *vosl* of the individual respondents can next be used to compute upper and lower bounds on the cumulative distribution of the *vosl* in the group of respondents. The results of this exercise are shown in Figure 3, as the upper and lower bounds on the cumulative distribution of

vosl implied by the consistent choices.[6] Again, we only show the relevant diagram for group 4; the other four diagrams are available upon request.

Figure 3. Upper and lower bounds on the cumulative distribution of the vosl, group 4 (vosl in mln. Dfl)



The real distribution of the *vosl* is unknown, since the ten choices made by the respondents reveal only an interval in which the *vosl* must be (under the assumptions made). There are three points in the diagram, for three values of the *vosl*, for which the lower and upper bounds coincide. These points correspond with the vertical lines in Figure 2. The questions associated with these vertical lines ask a respondent to indicate whether his *vosl* is higher or lower than a particular value. This provides exact information about the associated point on the cumulative distribution function of the *vosl*.

5 Extension towards a statistical model

The analysis of the previous section was based on two assumptions: (1) the *vosl* and *vot* are individual-specific constants and (2) each choice of a respondent reveals her true preferences. The second assumption is violated by respondents who made a sequence of choices that are not

[6] The upper and lower bounds referring to an estimated model will be discussed in the next section.

mutually consistent. This means that more than one third of the respondents had to be left out of consideration, which is clearly unsatisfactory. It would be preferable to have an extension of the half-space method that enables us to explain the occurrence of consistent and inconsistent choice sequences, and exploit the information provided by both. We will now provide such an extension by introducing an error generating mechanism into the hypothetical choice process. More specifically, we now assume that in every choice there is a fixed probability, q, that the respondent's choice is in accordance with his preferences. The probability $1-q$ that the choice differs from the preferences results from errors, caused for instance by pressing the wrong button. The situation considered in the previous two section corresponds to the special, deterministic case in which $q=1$.

The use of a fixed probability has the disadvantage that it assigns the same probability of 'erring' for a choice between alternatives that are nearly equivalent in term of utility for the respondent, as for a choice between alternatives where on is clearly preferable to the respondent. Conventional discrete choice models, in contrast, imply a smaller probability of choosing the alternative with the lower systematic indirect utility, when the difference in systematic utility between alternatives becomes bigger. The essential advantage of doing it this way, is that we can still avoid making *a priori* assumptions on the type of statistical distributions that apply.

To illustrate the extended model, we add the assumption that all respondents have a *vosl* and a *vot* that are both nonnegative. This implies that all respondents can be located somewhere in the nonnegative orthant of the *vosl*-*vot* space. Their true *vosl*-*vot* combination lies in that particular intersection that is defined by consistent choices that correspond with this combination. Let $m(k)$ be an index that denotes the particular intersection of half spaces to which the *vosl*-*vot* combination of a respondent of group k belongs; hence, in Figure 2 for $k=4$, each area gets a specific value of $m(k)$. (Because the lines of indifference vary between groups, we use group-specific indices $m(k)$).

If all respondents would only give answers that are in accordance with their true *vosl* and *vot*, the numbers indicated in the Figure 2 would imply the fractions of the respondents with *vosl*-*vot* combinations as implied by the boundaries of the intersection. However, when the respondents make errors, this is no longer the case. But we can then still estimate these fractions, on the basis of our assumption about the error-generating mechanism (the fixed probability q).

To do this, we assume that there is a joint density of *vosl* and *vot* in the population from which our respondents originate. We regard the actual *vosl*-*vot* combination of a respondent as a random draw from this density and denote the probability that a respondent in group k has a true combination of *vosl* and *vot* that belongs to a particular area $m(k)$ as $p_{,m(k)}$. We should have $p_{m(k)} > 0$ for all $m(k)$ and $\sum_m p_{m(k)} = 1$ for all groups *k*. We know each respondent's group *k*, but not the area to which his or her particular *vosl*-*vot* combination belongs. Since there is a unique set of choices associated with each area *m*, the area to which a respondent belongs is revealed by his choices if no mistakes are made. The probabilities p_{km} can then be estimated as the relative frequencies of respondents ending up in area *m*. This is the approach followed in the previous section. When respondents make errors, these probabilities can still be estimated, although the procedure is somewhat more complicated.

To see this, observe that the probability that a respondent makes a sequence of choices that is completely consistent with his *vosl* and *vot* is, with 10 choices, q^{10}. The probability that he makes exactly one choice that is inconsistent with his true preferences equals $(1-q)\cdot q^9$.[7] More generally, the probability that a particular sequence of choices – say y_i - is made by respondent *i* in group *k* whose true *vosl*-*vot* combination belongs to $m(k)$ equals $q^{n_{imk}}(1-q)^{(10-n_{imk})}$, where n_{imk} is the number of choices in y_i that are consistent with a *vosl*-*vot* combination in area *m*. This is the probability that y_i will be observed, conditional on the respondent's *vosl*-*vot* combination belonging to $m(k)$. The unconditional probability that y_i will be observed can be found by multiplying the conditional probability by $p_{k(m)}$ and summing over all $m(k)$s. The resulting expression can be interpreted as the likelihood of observing y_i for a respondent in group *k*:

$$l(\mathbf{y}_i, k) = \sum_m p_{m(k)} \cdot q^{n_{imk}} \cdot (1-q)^{(10-n_{imk})}. \qquad (5)$$

If no mistakes are made (q=1), equation (5) says that the likelihood of observing a particular sequence of choices is equal to the probability p_{km} that the respondent's *vosl*-*vot* combination belongs to the unique area *m* that associated with this choice sequence. When

[7] It may be noted that such a mistake may result in a sequence of choices that is consistent with the preferences of a respondent whose combination of *vosl* and *vot* belongs to another area than *m*. It is therefore possible that an observed sequence of choices is internally consistent, but nevertheless not in accordance with the preferences of the actor.

mistakes can be made (q<1) any area $m(k)$ can lead to any particular combination of choices, although of course not all combinations have the same probability.

We have estimated the probabilities $p_{k(m)}$ and q by maximizing the product of the likelihoods $l(y_i k)$ over all individuals in the same group k. Doing this for all groups gives us estimates for all $p_{k(m)}$s and four (potentially different) estimates for the probability q that a respondent's choice agrees with his actual *vosl*-*vot* combination.

Table 3 shows the five estimates for the probabilities q of making a choice in accordance with one's true preferences. The low standard errors show that these probabilities are estimated precisely. For all three groups these probabilities are between 90 and 95%, which seems reasonable. Note, that the probability of realizing a sequence of 10 correct choices is 35% when q=.90 and 60% when q=.95. This suggests that a non-negligible share of the consistent choices may in fact be inconsistent with the true preferences of the respondent. Note also that the five point estimates of the qs are very close to each other, which suggests that the five groups are equal in their propensity to make errors, as one should expect on the basis of the random assignment of respondents to groups.

Table 3 Estimated probabilities that a choice agrees with the respondent's preferences

Group	*q*	*standard error*	*Loglik*
1	0.930	0.0066	-882.64
2	0.943	0.0054	-870.60
3	0.932	0.0060	-899.20
4	0.936	0.0063	-954.93
5	0.944	0.0055	-801.69

Note. The probabilities $p_{m(k)}$ have been estimated jointly with the qs.

Using the estimation results for p_{km} (not reported, but available on request), we can compute alternative upper and lower bounds of the distribution function of the *vosl* in the same way as we did this in the previous section for the respondents with consistent choices. The estimation results imply somewhat different bounds for the cumulative distribution of the *vosl*. These are pictured in Figure 3 as the 'thin' lines. The bounds that are based on the statistical model have a tendency to lie above those based on consistent respondents only. This was the case

not only for group 4, but for each of the five groups However, the bounds are in general close to each other.[8]

An advantage of the model-based bounds is that they take into account the information contained in choices from respondents with inconsistencies, whereas the other bounds are based only on the consistent respondents. This advantage comes at the 'cost' of having to make assumptions about the way respondents make errors.[9]

Figure 3 shows that the stated choice experiment, when interpreted with the two or three (if one wants to incorporate the inconsistent choices) minimal assumptions used here is certainly informative. It contains valuable information about the distribution of the *vosl* among the respondents, and – provided these respondents are a-selectively chosen – in the population. Our assumptions are, however, insufficient to exactly identify this distribution, even though some points of the cumulative distribution are revealed exactly, when the lower and upper bounds touch, as happens three times in Figure 3.

Table 4 Upper and lower bounds for the median vosl.

	Consistent choices		Estimated model	
Group	lower	Upper	Lower	Upper
1	1.5	5.5	1.5	5.5
2	2.7	6.6	2.7	6.6
3	0.0	4.4	0.0	8.3
4	5.5	6.6	1.7	5.5
5	3.6	5.5	3.6	5.5
All	2.7	6.6	2.7	5.5

Millions of Dutch guilders.

To get a better idea of the sharpness of our results, Table 4 lists the implied upper and lower bounds for the median *vosl* in the five groups of respondents.[10] It is apparent from this Table that

[8] It may be noted that these bounds are based on the estimated values of the p_{km} and are therefore subject to estimation error.

[9] We already observed that choices that are internally consistent do not necessarily reveal true preferences, since errors do not necessarily result in an inconsistent choice sequence. It is therefore possible that an inconsistent choice sequence contains as much information about the true preferences of a respondent as a consistent sequence.

[10] Figure 3 gives the corresponding diagram.

the relevant intervals are typically very wide. The smallest one is 1.1 million Dutch guilders, the widest 8.3, and the average width is 3.85. The reason for this result is clear: given the questions posed to the respondents, and given our unwillingness to make further assumptions, it is simply impossible to reach more precise results with respect to the relevant intervals.

This suggests two possibilities two improve this result. One is to introduce more assumptions. This will be discussed in the next section. The other, relevant only in the design stage of a stated choice experiment, is to change the choice cards shown to the respondents. The exposition above has made clear that choice situation in which the two alternatives differ only in the number of fatalities and the toll reveal one point of the cumulative distribution function exactly (assuming consistent choices). More in general, the cumulative probability distributions will be closer when the minimum and maximum *vosl*'s consistent with each area become closer. Looking at Figures 1 and 2, and realizing that the positions of the lines of indifference follow from the differences in attribute levels between the alternatives in a choice set, it is clear that the closeness of upper and lower bounds, and hence the potential usefulness of the half-space method, can be affected through the underlying design of the experiment. If the interest of the survey is especially in the determination of the median *vosl*, the choice sets could be constructed in such a way that the lines of indifference are relatively close and relatively vertical in that part of the *vosl*-*vot* space where the median is expected to be.

Finally, it may be noted that more precise information about the cumulative distribution of the *vosl* can be collected by posing more questions to the respondents, but that it is at the same time likely that then also the frequency of inconsistent choice sequences will increase. Indeed, the analysis of the present section shows that even with modest probabilities of making an error, a substantial share of the choice sequences is not internally consistent and could therefore not be used in a deterministic model. Hence, the relevance of the statistical model increases with the size of the choice sequences.

6 Comparison with mixed logit models

The foregoing discussion implies that the half-space method cannot pin down a single distribution function of the *vosl*, but produces lower and upper bounds instead. Even the median value of the *vosl* is not precisely indicated by the results of the stated choice experiment.

Figure 4 Upper and lower bounds on the vosl distribution using all respondents' choices (vosl in mln Dfl).

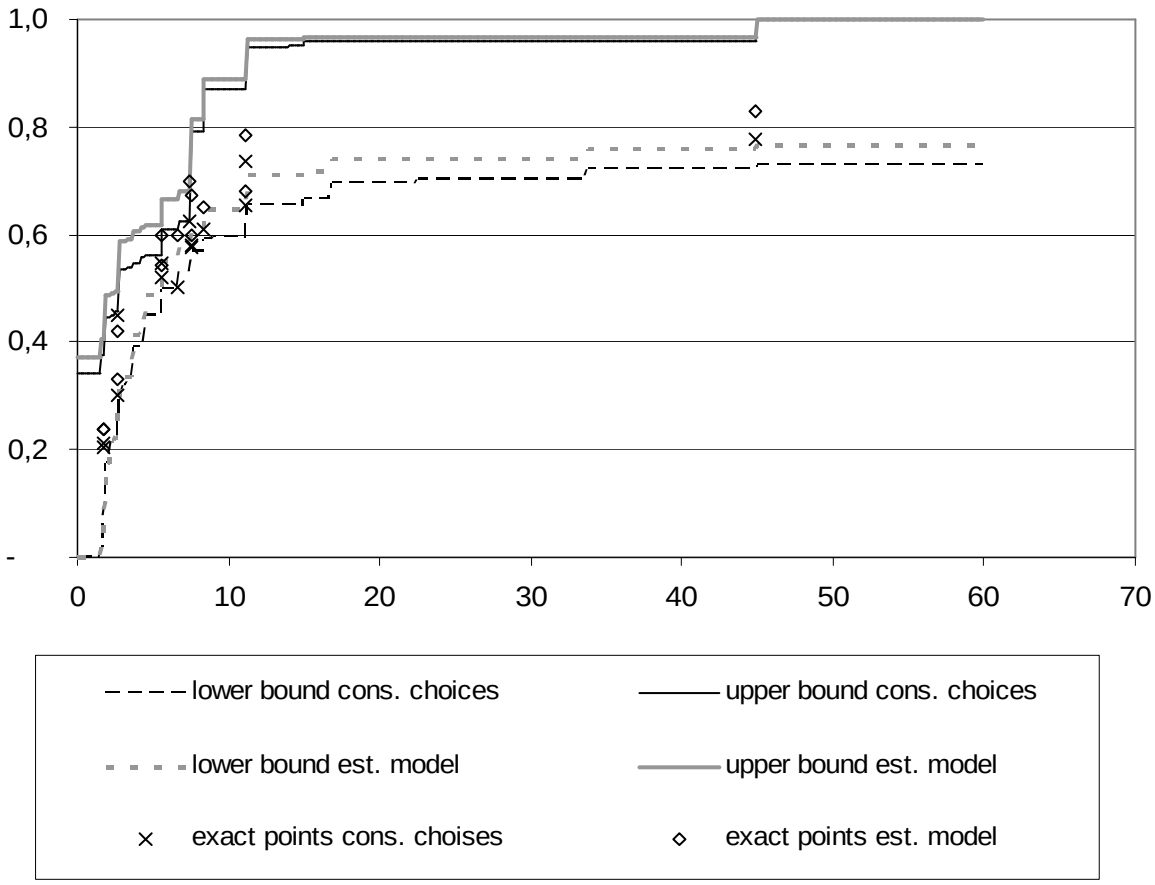


Figure 4 summarizes the information about the *vosl* in our complete sample, for all groups jointly (Figure 3 referred to group 4 alone). Figure 4 was obtained in the same way as Figure 3, but it uses the information of respondents from all five groups (the five blocks from the design). The diagram also pictures the points of the *vosl* distribution of the various groups that were exactly revealed by their choices (that is, the points where the upper and lower bounds coincide in the diagrams such as Figure 3). It is clear from the diagrams that those points are inconsistent across groups: for some values of the *vosl*, different values of the cumulative distribution function were exactly revealed. Nevertheless, these differences in choice behavior between the five groups are limited. For instance, if we estimate a standard logit model and allow the parameters to differ between groups, we find that the differences between estimated parameters are numerically small

and most of them are insignificant.[11] A striking feature of Figure 4 is that for a *vosl* equal to 44.9 million Dutch guilders, there is an exactly revealed point in the distribution function of group 1 that is close to the lower bound, suggesting that the true distribution of *vosl* may have a relatively fat tail.

Conventional discrete choice models produce a unique distribution of the *vosl* by making other assumptions. This section compares the results of such models with the results from the half-space method, summarized in Figure 4. We limit our attention to discrete choice models with a linear utility function; which is, for the data at hand, specified as in equation (2). The multinomial logit model assumes that the coefficients β_i are constants or deterministic functions of observed characteristics of the respondents. In mixed logit models, these parameters are allowed to be random variables, but their distribution has to be specified *a priori*. Both approaches are more restrictive than the one used above, which treats ratios between them as individual specific constants without assuming anything about their distribution.

Furthermore, logit models add a random term, ε, to the utility function for each choice alternative. These terms can be interpreted in various ways. Important possibilities are unobserved heterogeneity of the choice alternatives or respondents, specification errors, and a random element in choice behavior due to errors. In our study, respondents were instructed to imagine that the two roads among which they had to choose were identical in all respects, except for the three attributes toll, safety and travel time. Unobserved heterogeneity of the alternatives should therefore be no source of randomness, or a minor issue at most. Sources of randomness could then be differences between individuals, specification errors, and erratic responses.

The latter interpretation brings us close to the model of the previous section. It must, however, be pointed out that the error mechanism used there is different from the one that corresponds to this interpretation of the logit model. In the previous section we assumed a given probability that the choice indicated by the respondent is not in agreement with his actual preferences. This means that the probability of an error is independent of the respondent's evaluation of the two alternatives. If we interpret the random term in the logit model as an evaluation error, the probability that a respondent's choice is not in accordance with his or her preferences depends on the evaluation of the choice alternatives. If the utilities of the two

[11] If group1 is used as the reference group, all 12 parameters that measure differences with other groups are small and 9 of them insignificant (asymptotic t-value less than 1.96).

alternatives differ widely, the probability that an error will be made is then much smaller than when they are close. To see this, recall that the probability P_1 that a respondent chooses alternative 1 from a choice set consisting of two alternatives is:

$$P_1 = \frac{e^{u_1 - u_2}}{1 + e^{u_1 - u_2}} \tag{6}$$

where u_i is the evaluation of alternative *i* by the linear utility function. If $u_1 > u_2$, the choice for alternative 1 is in accordance with the respondent's preferences. This probability exceeds 0.5, and increases in $u_1 - u_2$. The probability of an erroneous choice is, accordingly, at most equal to 0.5 and it decreases in $u_1 - u_2$. This difference in the specification of the error mechanism is an important difference between the logit models and the approach to stated choice data of the previous sections of the present paper.

Since the two error generating mechanisms are different, we will in what follows compare the results of estimating logit models with respect to the distribution of the *vosl* with the bounds computed on the basis of the consistent choice sequences computed earlier in this paper.[12]

We have estimated a number of mixed logit models that differed in the specification of the distribution of the random coefficients.[13] We compared the implied distribution of the *vosl* with the bounds implied by the approach of this paper. Since the logit models are estimated on all respondents, we compare the implied distribution of the *vosl* with the bounds of the distribution of the *vosl* for all respondents as shown in Figure 4. Our illustration concentrates on mixed logit models in which the three coefficients are assumed to be lognormal distributed. We estimated these models in preference space as well as in willingness-to-pay space. In the former case we specify the utility of alternative *i* as:

$$u_i = \beta_{toll} \cdot \tau_i + \beta_{prob} \cdot P_i + \beta_{time} \cdot T_i) + \varepsilon_i \tag{7}$$

where the β parameters are assumed to be negative and lognormally distributed, and the error term ε_i extreme value type I distributed. Each individual respondent evaluates all alternatives on

[12] Since the bounds we derived on the basis of the consistent choice sequences are very close to those derived on the basis of the model with the error mechanism, this is not of much practical interest.

[13] Since Figures 2 and 4 strongly suggest the presence of heterogeneity in the *vosl*, no comparison with multinomial logit has been made.

Table 5 Estimation results for mixed logit models.

a) Estimation in preference space

			No correlation		*Free correlation*	
			Coeff.	*St.e.*	*Coeff.*	*St.e*
Toll	β_1	μ_1	-0.44	*0.044*	-0.39	*0.059*
		σ_1	0.94	*0.040*	1.09	*0.072*
# Fatalities	β_2	μ_1	-1.92	*0.058*	-1.78	*0.068*
		σ_1	1.30	*0.058*	1.37	*0.081*
Travel time	β_3	μ_1	-2.01	*0.070*	-2.15	*0.112*
		σ_1	0.85	*0.076*	1.07	*0.069*
		η_{12}			-0.05	*0.10*
		η_{13}			0.83	*0.13*
		η_{23}			-0.07	*0.11*
Loglikelihood			*-4512.87*		*-4485.58*	

b) Estimation in wtp space

			No correlation		*Free correlation*	
			Coeff.	*St.e.*	*Coeff.*	*St.e*
Scale factor	β_1'	μ_1	-0.15	*0.061*	-0.34	*0.059*
		σ_1	1.10	*0.078*	1.08	*0.073*
VOSL	β_2'	μ_1	-1.54	*0.045*	-1.42	*0.066*
		σ_1	1.80	*0.064*	1.48	*0.079*
VOT	β_3'	μ_1	-1.87	*0.068*	-1.75	*0.10*
		σ_1	1.05	*0.047*	1.00	*0.073*
		η_{12}			-0.97	*0.092*
		η_{13}			-0.28	*0.10*
		η_{23}			-0.069	*0.11*
Loglikelihood			*-4531.91*		*-4483.13*	

the basis of the same realization of the coefficients β. The error terms of all alternatives are assumed to be independent of each other.

Our second specification estimates the model in willingness-to-pay (*wtp*) space. In that case the utility of alternative *i* is specified as:

$$u_i' = \beta'_{toll} \cdot \left[\tau_i + \beta'_{prob} \cdot P_i + \beta'_{time} \cdot T_i \right] + \varepsilon_i \,. \qquad (8)$$

The β' parameters are again assumed to be lognormally distributed, β'_{toll} negative and β'_{prob} and β'_{time} positive, and the error term ε_i extreme value type I distributed. Note that (7) is identical to (8) if we have $\beta_{toll} = \beta'_{toll}$, $\beta'_{prob} = \beta_{prob} / \beta'_{toll}$ and $\beta'_{time} = \beta_{time} / \beta'_{toll}$ It is referred to as a specification in willingness-to-pay space because β'_{prob} and β'_{time} are the *vosl* and the *vot*, respectively. These *wtp* variables are therefore immediate results of the model estimation if (8) is used, whereas they have to be computed on the basis of estimates of the estimated βs if (7) is used. With either model, the distribution of the *vosl* and the *vot* is lognormal, and one would expect that the two specifications give similar results.

However, using different data, Sonnier, Ainslie and Otter (2005) and Train and Weeks (2005) have reported the somewhat surprising finding that a better within-sample fit was obtained for the model estimated in preference space, but 'more reasonable' *wtp* distributions for the model estimated in *wtp* space.[14] It appears, therefore, that very small differences in model specification may give rise to substantial differences in results, and the authors suggest that one method should be preferred to the other as being more reliable. Our approach allows a specific check on the plausibility of the *wtp* distributions, by comparing the outcomes of both models with the distributions generated by the half-space method.

To do so we estimated the mixed logit models described above in preference space as well as in willingness-to-pay space. For both specifications, two variants were estimated: one in which the three coefficients were treated as independent lognormal variables, and another in which they were assumed to be simultaneously distributed with unrestricted correlation parameters. Estimation results are given in Table 5. They are satisfactory for both specifications. When no correlation between the βs is allowed, the model estimated in preference space has a larger

[14] See Train and Weeks (2005) page 16. Train and Weeks refer to an earlier version of Sonnier et al. (2005). Unlike what was the case in these studies, we have no objective measure for what would be a reasonable or unreasonable willingness to pay.

loglikelihood value at convergence, as was found in the two papers discussed above, but this difference disappears in the more general model.

Figure 5 The vosl distribution implied by the mixed logit models with lognormal coefficients (preference space vs. willingness to pay space).

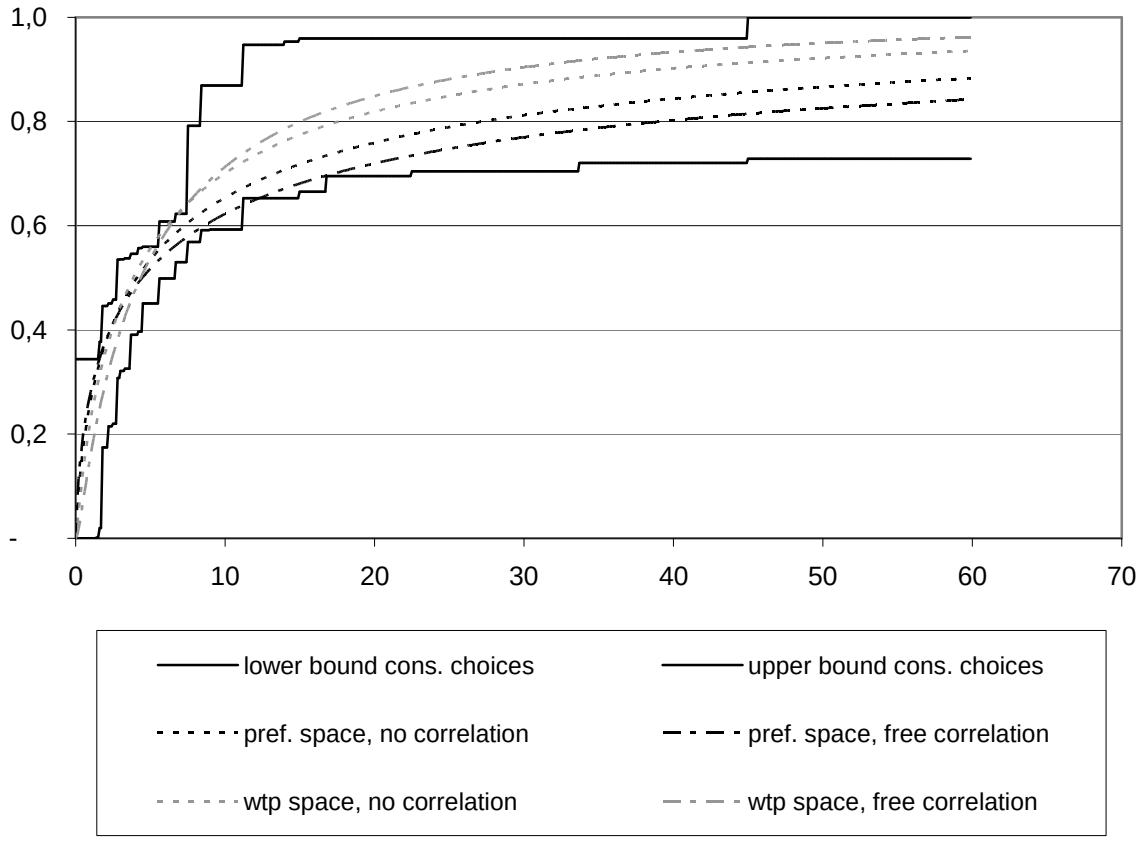
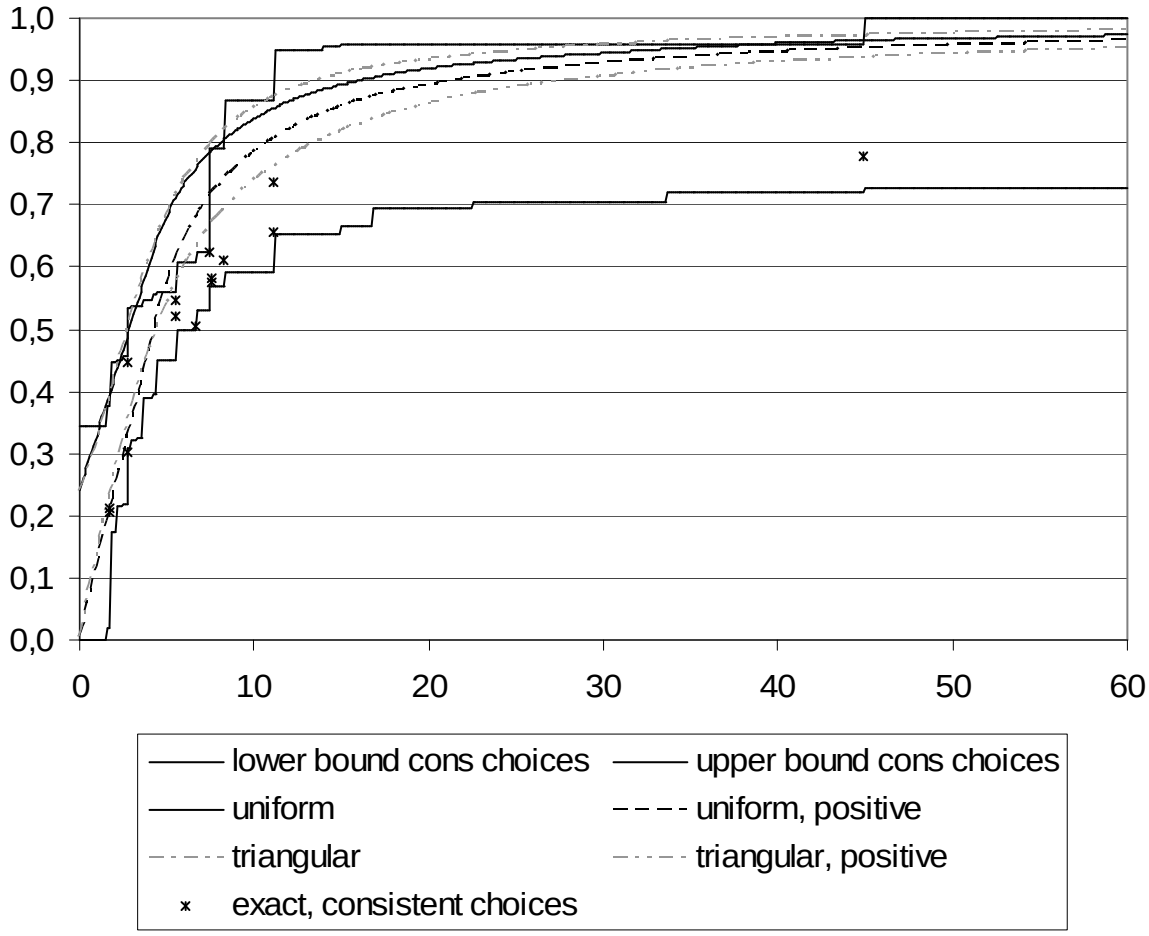


Figure 5 shows the distributions of the *vosl* implied by these four mixed logit models, and compares them with the bounds implied by the consistent choice sequences. The picture shows that, except for low values of the *vosl*, the curves resulting from estimation in preference space are closer to the lower bound than those from estimation in willingness to pay space. This replicates the results obtained by Sonnier et al. (2005) and Train and Weeks (2005). However, the bounds derived earlier in this paper show that the information content of our sample does not

allow us to regard one of the two specification types as giving more reasonable results than the other. It appears that the specification in preference space implies a *vosl* distribution that is closer to the lower bound derived with the half-space method. Even though one could, for *a priori* reasons, regard the tail of the *vosl* distribution implied by the preference space specification as unreasonably fat, it should be noted that the lower bound derived earlier in this paper is consistent with an even 'fatter' tail. Moreover, the single point of the distribution function for higher values of the *vosl* that was exactly revealed by group 1, shown in Figure 4 at a *vosl* of 44.9, suggests that the lower bound we computed might be closer to the true distribution function than the upper bound.

Figure 6 The vosl distribution implied by the mixed logit models with uniform and triangular coefficients.



This observation also suggests that it may be not always be a good strategy to impose probability distributions on the random coefficients that are bounded on both sides. Such bounded distributions of partworths have been discussed, for instance, in Train and Sonnier (2003). To investigate this issue, we estimated a number of logit models with uniform and triangular distributions. In both cases we found that a substantial part of the probability mass was assigned to positive values of the coefficients. Since there is nothing in our data that suggests that some of our respondents attach positive value to tolls, unsafety or travel time, we also estimated versions of the models in which the random coefficients were restricted to be nonnegative.

Figure 6 shows the implied distributions of these mixed logit models and compares them to the bounds we computed with the half-space method for the consistent choices.[15] The diagram shows that the distributions from these mixed logit models are indeed close to the upper bounds form the half-space method, as was expected. Moreover, all estimated distributions cross the upper bound for values of the *vosl* just above the median. This is not only the case for the upper bound that refers to consistent choices, but also for the bounds computed on the basis of the model we estimated in section 5 (not shown in Figure 6), which is just a little bit higher than the one referring to consistent choices (compare Figure 4). Moreover, all these curves are above most of the points of the distribution functions of the *vosl* that were revealed exactly by the choices of the various groups. This suggests that the use of bounded partworths may underestimate the variation in the parameters of interest that is present in the data.

7 Results for the value of time

We carried out a similar analysis for the value of time (*vot*) on the same data. The results are summarized in Figures 7 and 8, which can be compared to Figures 5 and 6. The lower and upper bounds computed for the *vot* differ more (in a relative sense) than those computed for the *vosl*.

Figure 7 shows the distributions of the *vot* implied by the mixed logit models with lognormal coefficients that have been discussed in the previous paragraph. It shows the same pattern as Figure 5: the models estimated in *wtp* space imply a distribution of the *vot* that is much closer to the upper bound than those based estimated in preference space.

[15] Full estimation results are available on request.

Figure 7 The vot distribution implied by the mixed logit models with lognormal coefficients (preference space vs. willingness to pay space).

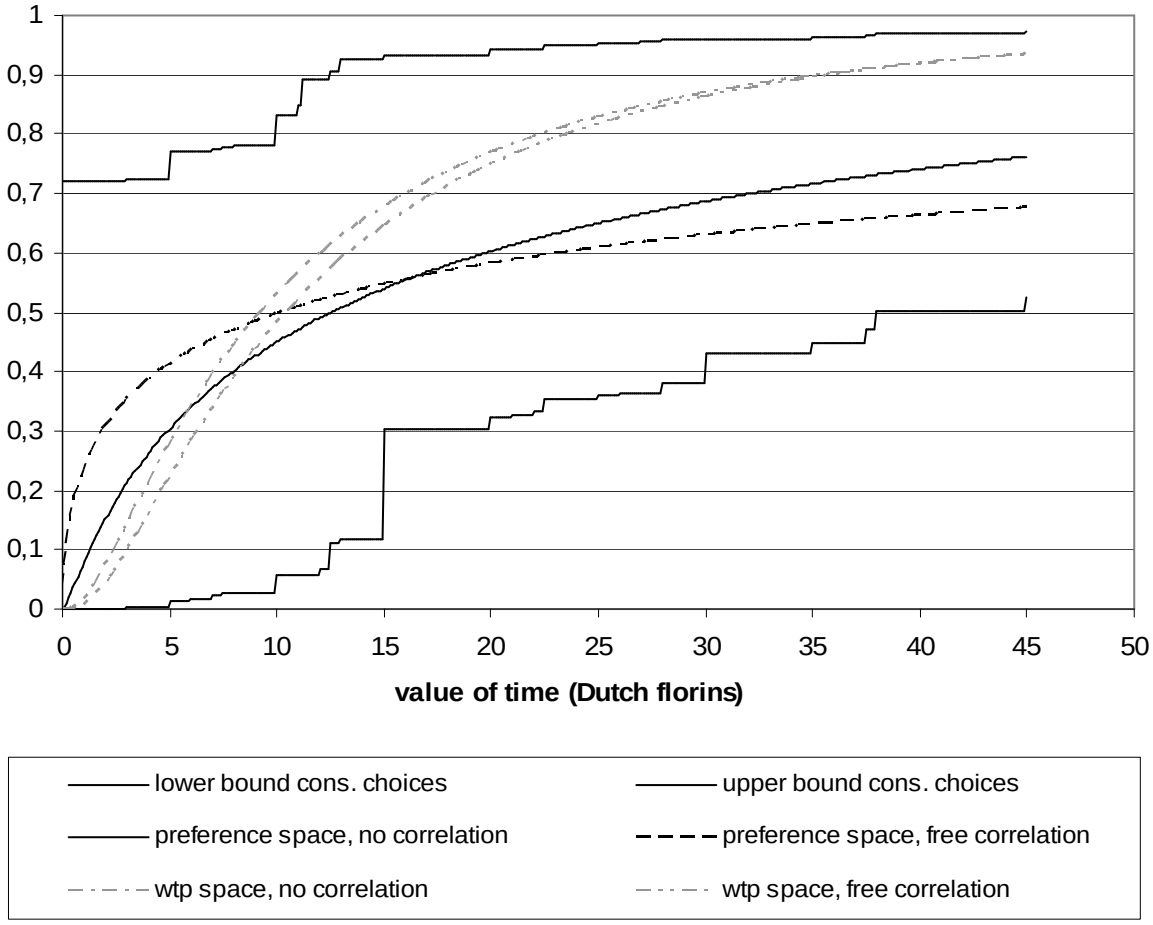
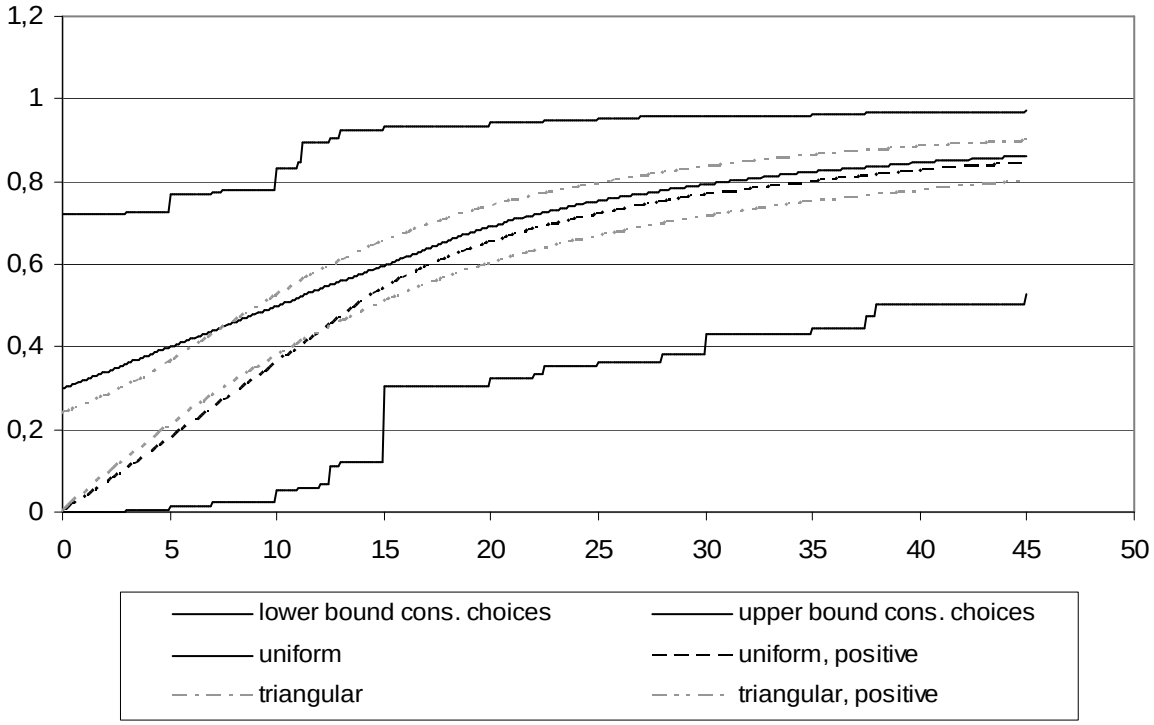


Figure 8 shows the estimated distributions for the *vot* when the uniform and triangular distributions are used. Also in this case, restricting the random coefficients to be positive implies that the whole distribution shifts upward. Even though none of the models now implies a distribution that crosses the upper bound, there appear to be no reason to be prefer the models with bounded partworths to the lognormal models. Also in this respect the results for the *vot* confirm those reached in the previous sections for the *vosl*.

Figure 8 The vot distribution implied by the mixed logit models with uniform and triangular coefficients (vot in Dfl).



8 Potential for and limitations of the half-space method

The foregoing sections illustrated how the half-space method can be used to determine objectively upper and lower bounds on the (cumulative) distributions of *wtp*'s (*vosl* and *vot* in this application) from SP experiments. It is clear that there is an advantage in deriving these bounds directly from the actual responses, instead of making *a priori* assumptions on the type of distribution. It has also become clear that the method will become more effective when more choices are included that imply an equality of the lower and upper bounds of the distributions, thus narrowing the distance between the bounds. These are choices between alternatives for which only the price attribute and the attribute to be valued differ. Evidently, other considerations in the design phase of an SP experiment may lead to a design in which such choices are *not* included, so the ultimate design would in part depend on the question of how important a direct determination of the type of distribution is considered to be.

The method is not restricted to studies with binary choices. For example, a choice between three alternatives can be treated as two binary choices, namely between the winning alternative and either of the two unchosen alternative.

Our application has also illustrated a number of difficulties with this half-space method. One difficulty is the inconsistency of respondents, causing a "true" combination of *wtp*'s that explains all their choices to be non-existent. A rather rigorous approach of considering "consistent" respondents only was found to still produce usable distributions, but is of course unattractive due to the loss of valuable data – an loss that will become bigger, relatively, when the number of choices per respondent increases. We proposed a somewhat naïve, *ad hoc* but intuitive error-generating mechanism that allowed us to keep the "inconsistent" respondents. It seems that this mechanism is still amenable to improvement, for example by somehow making the probability of erring larger for a choice that seems to deviate more from the respondent's other choices. But we admit that inconsistencies of this type limit the attractiveness of the proposed method, and ways to deal with it will probably always remain somewhat unattractive.[16]

A second possible limitation is related to the dimensionality of the design. Our application considered two attributes besides price, allowing for a two-dimensional graphical representation. If three *wtp*'s or more are at stake, the essence of the method will not change, but the relations between the *wtp*'s will become more complicated than what is shown in equation (4) and the interpretation of the sub-spaces becomes more difficult. The method, therefore, seems to be best suited for one- or two-dimensional *wtp* studies.

Third, when different "blocks" are used, as in our study, it is not sufficient to have for each group choices that imply an equality of the lower and upper bounds of the distribution to achieve also an "exact" point for the entire population. Figure 4 gives an example. To circumvent this problem, it would suffice to use identical choices of this type across groups. One such question near the expected median presumably would have reduced the gap between the lower and upper bounds shown in Figure 4 considerably. Another one, near the end of the range, would have given valuable information on the likely "fatness" of the tail.

[16] A reviewer remarked that the alternative possibility of allowing for non-linearities in the utility function is also not entirely satisfactory. For example: which non-linearities and interactions should be included? And should these be included also for "consistent" respondents?

9 Conclusion

In this paper we developed the half-space method for investigating the results of stated choice experiments under minimum assumptions on statistical distributions. The method provides upper and lower bounds for the distribution of willingness-to-pay variables that are often the focus of interest of such experiments, under the assumption that an individual's willingness to pay is constant in the range investigated. As a by-product, the method provides a check for the overall consistency of the sequences of choices made by the respondents.

The half-space method also sheds light on questions about the appropriate specification of the distribution of the random coefficients in mixed logit models. We showed that, for the data at hand, the method implies a relatively broad range between the upper and lower bounds of the distribution function of the *vosl*. Many parametric distributions of the *vosl* are within this range. Our comparison of models estimated in preference space and in willingness-to-pay space showed that details in the specification can lead to estimated distributions of the *vosl* that are either close to the upper or to the lower bound implied by our analysis. Since the data do not allow us to make a choice, our analysis suggests that it might be useful in the design phase of a stated choice experiment to consider inclusion of sufficient choices that imply an equality of the lower and upper bounds of the *wtp* distributions determined by the half-space method (thus narrowing the distance between the bounds). These are choices between alternatives for which only the price attribute and the attribute to be valued differ.

Incorporation of such choice cards in an SP design , preferably near the expected median and in the tail of the distribution to allow a check on "fatness", seem particularly relevant when it is expected that the *wtp*'s to be investigated may vary strongly over respondents, so that mixed logit models become a likely option for the analysis, and direct information on the distribution over respondents becomes more valuable. Due the specific pitfalls of the half-space method, it will probably be impossible to develop it into a full alternative for conventional discrete choice models, but the method does provide a nice way of getting direct information, under minimal assumptions, on the distribution of *wtp*'s implied by the responses to an SP questionnaire. This seems to make it a very useful complement to mixed logit models, that require *a priori* assumptions on these distributions.

References

Ashenfelter, O. (2006) Measuring the Value of a Statistical Life: Problems and Prospects. *Economic Journal*, **116**, C10-C23.

Hensher, D.A., J.M. Rose and W.H. Greene (2005) *Applied Choice Analysis: A Primer*, Cambridge University Press.

Loomes, G., P. G. Moffat and R. Sugden (2002) A Microeconometric Test of Alternative Stochastic Theories of Risky Choice. *Journal of Risk and Uncertainty*, **24**, 103-130.

McFadden, D. and K. Train (2000) Mixed Multinomial Logit Models for Discrete Response. *Journal of Applied Econometrics*, **15**, 447-470.

Meijer, E. and J. Rouwendal (2006) Measuring Welfare Effects in Models with Random Coefficients. *Journal of Applied Econometrics*, **21**, 227-244.\

Rizzi L I, Ortúzar J de D, 2003Stated preference in the valuation of interurban road safety *Accident Analysis and Prevention*, **35**, 9 – 22.

Sonnier, G., A. Ainslie and T. Otter (2005) Estimating Willingness to Pay with Random Coefficient Choice Models. Working paper.

Train, K. and G. Sonnier (2005) Mixed Logit with Bounded Distributions of Partrworths. Chapter 7 in : A. Alberini and R. Sharpa (eds) *Applications of Simulation Methods in Environmental and resource Economics*, Springer.

Train, K. and M. Weeks (2005) Discrete Choice Models in Preference Space and Willingness-to-Pay Space. Ch. 1 in: A. Alberini and R. Sharpa (eds) *Applications of Simulation Methods in Environmental and resource Economics*, Springer.

Wedel, M., W. Kamakura, N. Aurora, A. Bemmaor, J. Chiang, T. Elrod, R. Johnson, P. Lenk, S. Neslin, C.S. Poulsen (1999) Discrete and Continuous Representations of Heterogeneity in Choice Modeling. *Marketing Letters*, **10**, 219-232.

Wesemann, P., A.T. de Blaeij en P. Rietveld (2005) De waardering van bespaarde verkeersdoden (The valuation of saved lives in transport; in Dutch). SWOV report.

Appendix A Stated choice question

Suppose that you would like to travel from city Y to city Z for personal reasons. Assume that you will be traveling alone by car. You can choose between two roads. Both roads are used equally intensive with 55000 trips per day, which means around 18 million trips a year. So in one year, every Dutch citizen could have used this road.
Both roads are toll roads, and you have to pay the toll yourself. You have to make a choice between the roads based on 3 criteria; safety, travel time and toll. The roads are otherwise identical: they are equally beautiful, calm, interesting, etc. So the only possible differences between the roads are the tolls, the travel times and traffic safety. All the other characteristics are equal and should not play a role in your decision making.

If the choice were between road A and B, which road would you choose?

Choice	Road A	Road B	Choice
Toll	ƒ 5,-	ƒ 10,-	❑ Road A
Number of fatal accidents a year	10	5	❑ Road B
Travel time	1 hour	50 min	