Transportation Research Part B 44 (2010) 136–151

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

The information content of a stated choice experiment: A new method and its application to the value of a statistical life

Jan Rouwendal [a,*], Arianne de Blaeij [b], Piet Rietveld [a], Erik Verhoef [a]

[a] Department of Spatial Economics, VU University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
[b] LEI, Public Issues Division, P.O. Box 29703, 2502 LS Den Haag, The Netherlands

ARTICLE INFO

Article history:
Received 19 December 2006
Received in revised form 24 April 2009
Accepted 24 April 2009

Keywords:
Stated preferences
Value of a statistical life

ABSTRACT

This paper presents a new method to assess the distribution of values of time, and values of statistical life, over participants to a stated choice experiment. The method 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. It 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.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Discrete choice analysis, of both stated and revealed preference data, is an important method in 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

* Corresponding author. Tel.: +31 20 5986093; fax: +31 20 5986004.
E-mail address: jrouwendal@feweb.vu.nl (J. Rouwendal).

0191-2615/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trb.2009.04.006

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 Russell, 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).

In short, applications of mixed logit models suffer from what Manski (2007, p. 2) calls the law of decreasing credibility: 'The credibility of inference decreases with the strength of the assumptions maintained.' The problem is not specific to stated preference techniques or mixed logit models. Researchers often find themselves in situations in which the relevant theory does not suggest a fully parameterized model. To avoid arbitrary assumptions, flexible parametric and semi- and nonparametric techniques have been developed.[1] For instance, Fosgerau and Bierlaire (2007) use flexible polynomials to test the appropriateness of a particular choice of the mixing distribution in mixed logit models, and Rigby et al. (2009), using a Bayesian approach, propose the use of the Marginal Likelihood to compare model performance under various distributional assumptions.

Since any choice of a parametric mixing distribution is arbitrary, it is clearly 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*), that allows investigation of the mixing distribution while avoiding any *a priori* assumption about its functional form. This paper proposes such a method for exploring the distribution over individuals of the *vosl* and *vot*, or similar variables. The minimal *a priori* assumptions on which the method is based 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 min 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 Euros 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.[2] 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.

The closer these bounds are, of course, the more informative they are of the distribution of point estimates of marginal valuations. 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.

The fact that our method does not fully identify the mixing distribution is a consequence of our wish to avoid stronger *a priori* assumptions than the two mentioned above. In this respect it is not unique. There are many other situations in which the wish to avoid questionable assumptions implies that point identification is impossible, although it is still possible to restrict the variable of interest to intervals, that can be quite informative for the phenomenon under study. See, for instance, Blundell et al. (2007) for a study of changes in the distribution of wages and the references cited there.

We discuss the method of the present paper using data that were collected with the prime objective to investigate the value of statistical life (*vosl*) in road traffic for Dutch citizens,[3] producing value of time (*vot*) estimates as an intended by-product. The *vot* and *vosl* distributions derived show a 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

[1] See Pagan and Ullah (1999) for an overview of the field. This book includes a chapter about discrete choice models, including a discussion of the nonparametric method for estimating the binary choice model developed by Matzkin (1992). That paper, assumes – like most of the related literature - that choices are determined by the sum of a deterministic function of explanatory variables and a random variable. This excludes mixed logit models.

[2] Note that half-spaces in two-dimensional space are also sometimes referred to as half-planes. The word "half" does not imply that both half-spaces should be equally big.

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

Download English Version:

https://daneshyari.com/en/article/1132477

Download Persian Version:

https://daneshyari.com/article/1132477

Daneshyari.com