Transportation Research Part B 81 (2015) 289–301

Contents lists available at ScienceDirect

Transportation Research Part B

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

# Commuters' preferences for fast and reliable travel: A semi-parametric estimation approach ☆

Paul R. Koster *, Hans R.A. Koster

*Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*
*Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands*

ARTICLE INFO

*Article history:*
Received 29 May 2013
Received in revised form 29 April 2015
Accepted 16 May 2015
Available online 4 June 2015

*Keywords:*
Local maximum likelihood
Latent class
Semi-parametric logit
Normalised entropy criterion
Value of time
Value of reliability

ABSTRACT

We employ a semi-parametric estimation approach to analyse observed and unobserved heterogeneity in the value of savings in travel time and schedule delay. Our econometric approach allows for the estimation of unobserved and observed heterogeneity in preferences in a flexible way, meaning that we do not put any structure on how individual characteristics (such as income and age) relate to the value of savings in travel time and schedule delay. Using data from a stated choice experiment, we illustrate the estimation approach and find that there is substantial heterogeneity in the value of savings in travel time and schedule delay. For our data, we find that unobserved heterogeneity is more important than heterogeneity related to individual characteristics.



## 1. Introduction

Preference heterogeneity is of key importance when evaluating the effect of transport economic policies. For example, the success of off-peak rewarding policies is strongly related to the potential of travelling outside the peak for a group of travellers. Some studies have shown that the distributional effects of congestion pricing policies strongly depend on the heterogeneity in individuals preferences (Arnott et al., 1992; Arnott et al., 1994; Van den Berg and Verhoef, 2011a; Van den Berg and Verhoef, 2011b; Börjesson and Kristoffersson, 2014). Furthermore, heterogeneity is important when evaluating the benefits of private provision of highways, because operators' profits depend on the marginal willingness to pay (WTP) for reductions in travel time and travel time variability of travellers (Mills, 1981; Winston and Yan, 2011; Tan and Yang, 2012). Heterogeneity is also important when studying the effects of congestion pricing in the presence of alternative privately or publicly operated transport modes (Huang, 2000; Van den Berg and Verhoef, 2013). In all cases, using average WTP values may lead to biased welfare estimates and imprecise policy recommendations.

☆ Paul Koster gratefully acknowledges the financial support of the Advanced ERC Grant OPTION # 246969. Hans Koster acknowledges the financial support of Platform 31. We like to thank two referees for their very useful suggestions for improvements. Stefanie Peer, Yin-Yen Tseng, Jasper Knockaert and Erik Verhoef are thanked for their contribution to the development of the survey. Part of the work was done while the first author was visiting the Danish Technical University, Department of Transport. We like to thank Thijs Dekker, Vincent van den Berg, Mogens Fosgerau, Katrine Hjorth, Richard Paap, Piet Rietveld, Jan Rouwendal, Erik Verhoef and two anonymous referees, for their valuable comments. This paper has been presented at the ERSA conference in Palermo, and at the International Choice Modelling Conference 2011 in Oulton Hall, England. Seminar participants are thanked for their constructive comments and suggestions.

* Corresponding author at: Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Tel.: +31 20 5982369.

*E-mail address:* p.r.koster@vu.nl (P.R. Koster).

http://dx.doi.org/10.1016/j.trb.2015.05.011
0191-2615/

In this paper, we use a semi-parametric econometric framework to analyse heterogeneity related to observed individual characteristics, as well as unobserved heterogeneity, using panel data from a stated choice experiment. We estimate a semi-parametric panel latent class model that allows for a flexible specification of both unobserved and observed heterogeneity. Our method is applied to the estimation of value of travel time savings (VOT) and the values of savings in schedule delay early (VSDE) and late (VSDL). We obtain distributions of preferences, by employing local-likelihood estimation methods introduced by Tibshirani and Hastie (1987) and Fan et al. (1995, 1998), and by assuming that individuals who are similar in terms of characteristics (e.g. income, age) will have more similar values of savings in travel time and schedule delay.

Our semi-parametric panel latent class model has several features. First, our method does not make any assumptions on the shape of the distribution of preferences that is estimated. Second, our econometric procedure allows for the estimation of both unobserved and observed heterogeneity. It extends the local logit model used by Fosgerau (2007), Börjesson et al. (2012) and Hjorth and Fosgerau (2012) and the local multinomial and local cross-sectional latent class model of Dekker et al. (2014) to allow for unobserved heterogeneity and panel data. Third, kernel smoothing techniques reduce the well-known curse of dimensionality, meaning that the inclusion of a large number of covariates will not necessarily lead to imprecise estimates (but it may increase the bias compared to a fully saturated model). The use of kernel smoothing allows for the incorporation of all interactions between individual characteristics and preferences without the need to specify these interactions explicitly. Hence, our approach is particularly useful for datasets with a large number of individual characteristics. Fourth, our estimation method takes into account the repeated nature of the choices and therefore correctly deals with the panel dimension of the data. It results in a unique semi-parametric distribution of preferences for each unique combination of individual characteristics in our dataset.

The data to estimate commuters' VOT and value of arriving at the preferred arrival time at work are obtained from a stated choice survey held among participants of a real-world rewarding experiment to combat congestion. We estimate the VOT, VSDE and VSDL using the classical linear specification of schedule delay (Vickrey, 1969).

We find that the average values are somewhat higher than in the previous literature, most likely because respondents participate in a reward experiment and people are arguably more sensitive to costs than to rewards. We also show that there is substantial heterogeneity in the estimated VOT, VSDE and VSDL. The VOT, for example, ranges from about €20 to €80. The results suggest that unobserved heterogeneity seems to be more important than heterogeneity related to observable individual characteristics. Nevertheless, we find statistically significant impacts of income, education and household composition on the VOT, the VSDE and the VSDL.

The paper continues as follows. Section 2 discusses the design and set-up of the stated choice experiment and specification of the utility function, which is followed by the econometric framework in Section 3. Section 4 presents the main results and Section 5 concludes and discusses the limitations and directions for further research.

## 2. Stated choice experiment and utility specification

### 2.1. Stated choice experiment

We use data from a stated choice experiment among Dutch morning-commuters participating in a peak-avoidance reward experiment. In order to reduce congestion, these commuters received a reward if they did not travel between cameras A and B during the morning peak (6:30–9:30). These cameras were placed on a congested highway. An example of a choice scenario is given in Fig. 1. Respondents were asked to choose between two departure times. To account for travel time variability, each departure time has two possible travel times with a corresponding probability, arrival time at work and reward. Hence, commuters made a trade-off between earlier or later arrivals, shorter or longer travel times, and receiving a monetary reward. The preferred arrival time of the traveller was given as a reminder. It was based on previous questions in the questionnaire and was defined as the time that a traveller would like to arrive at work if there is no possibility to receive a reward, and there is no travel time delay. The lay-out was pre-tested in a focus-group, and internet pre-tests were carried out to ensure that respondents understand the questions well. Several considerations regarding the screen lay-out were made. First, we made explicit what is important for travellers: departure time, probabilities, travel times, arrival times and rewards. Second, travel times of the separate parts of the trip were given. This is to show the respondents why they (do not) receive the reward and how the travel time is built up from the different components. Fourth, we used bold values for the variables that would have been (potentially) important in the trade-offs of the respondents.

The attribute values for travel times were pivoted around the average travel time of the respondent to enhance realism (Hensher, 2001). Arrivals at Camera A were spread over the whole peak hour to have sufficient variation in arrival times. Several other constraints were put in the design attributes to enhance realism. These are described in detail in Knockaert et al. (2012). The experimental design was pre-tested using extensive simulation, so that we are sure that the design is able to recover a broad range of parameters.

We exclude respondents who chose randomly or for whom no individual characteristics are available (29% of the data). For the remaining respondents we have information on gross monthly household income, level of education, gender, age, household composition (single, children, etc.) and their residential and working location. The summary statistics of the