Available online at www.sciencedirect.com
ScienceDirect
Transportation Research Part E 44 (2008) 1025–1038
TRANSPORTATION RESEARCH PART E
www.elsevier.com/locate/tre

Joint RP–SP data in a mixed logit analysis of trip timing decisions

Maria Börjesson *

Royal Institute of Technology, SE-100 44 Stockholm, Sweden

Received 14 September 2006; received in revised form 9 October 2007; accepted 25 November 2007

Abstract

In the current paper, a departure time choice model including travel time variability is estimated, combining stated preference and revealed preference data. We account for response scale differences between RP and SP data and, applying the mixed logit model, test for correlation of scheduling sensitivity across RP and SP choices within individuals. The analysis implies systematic differences in the RP and SP data. With support of the evaluation from the Stockholm trial, this indicates that SP is less trustworthy for trip timing analysis and forecasting, presumably because there are temporal differences in RP and SP choice situations.

Keywords: Revealed preference; Stated preference; Mixed logit; Unobserved heterogeneity; Travel time uncertainty; Schedule delay

1. Introduction

Trip timing choice has received increased interest as a consequence of the congestion charging trial in Stockholm 2006. It has long been expected that driver's preferences for different departure times are an important dimension for policy evaluation in congested regions. In this paper, a departure time choice model for Stockholm is developed, taking into account that drivers face an uncertain travel time. Mode choice is considered partially, by modelling the propensity to switch from driving. It does not, however, take into account that public transport travellers may switch to driving. We further test for taste heterogeneity in scheduling delay and travel cost sensitivity, applying a mixed logit model structure.

Stated preference (SP) data have in the present case many advantages. The high correlation of mean travel time and travel time uncertainty in revealed preference (RP) data makes accurate estimation of the trade-offs unfeasible. Experience suggests that hypothetical choices are understood by the respondents reasonably well and that they are capable of making fairly accurate trade-offs between the attributes (Louviere et al., 2000). In particular, SP data should work relatively well for departure time choice because drivers have genuine

* Fax: +46 8 688 77 32.
E-mail address: maria.borjesson@wspgroup.se

1366-5545/$ - see front matter © 2008 Published by Elsevier Ltd.
doi:10.1016/j.tre.2007.11.001

experience of the involved attributes. There is, however, a general opinion that the overall response scale is distorted in hypothetical choice situations, which is a problem if the model is to be used for application. A tradition of combining RP and SP data, taking advantage of both, is therefore now a well established methodology in consumer behaviour analysis (Bhat and Castelar, 2002).

Combining RP and SP data have a long tradition in discrete choice modelling in particular (Ben-Akiva and Morikawa, 1990). The vast majority of earlier studies have applied MNL model structures. Bhat and Castelar (2002) and Hensher and Greene (2001) have, however, applied mixed logit structures to incorporate the effect of unobserved taste heterogeneity. Brownstone et al. (2000) applied a mixed logit model for automobile choice, to allow for heteroscedastcity and correlation across alternatives. Bhat and Castelar (2002) discuss the importance of unobserved response heterogeneity, scale differences, inter-alternative error correlations and state dependency effects in joint RP and SP estimation.

In a previous paper (Börjesson, 2006), a departure time and mode switch model was estimated using SP data only, and the present paper extends this analysis by using joint RP and SP data. No joint RP and SP analysis of trip timing choice has previously been published. The analysis therefore provides a unique opportunity to compare observed and stated trip timing behaviour. This study also provides an unique opportunity to study valuation of travel time variability in a joint RP and SP data framework.

Most choice models including travel time variability have used SP data (Black and Towriss, 1993; Bates et al., 2001; Noland et al., 1998; Small et al., 2000; Jackson and Tucker, 1981), although there is a general econometric tradition of favouring RP data. Lam and Small (2001) and Brownstone et al. (2005) are exceptions and have used RP data from a HOT-lane[1] project in California in order to estimate willingness to pay for reduction in travel time variability. They found, however, that valuation of travel time variability depend strongly on model specification, which calls for more research.

The scarcity of studies on travel time uncertainty using RP data depends primarily on two factors. First, as mentioned, there is normally high correlation between mean travel time and travel time uncertainty in RP data. Secondly, data on time-of-day variation in travel time uncertainty is seldom available with the required level of detail.

It is due to the evaluation of the Stockholm congestion charging trial, that RP data on travel time and travel time uncertainty have become available. The present paper uses this data in a joint analysis with SP data. RP and SP data was collected in Stockholm, April–May 2005. The respondents participating in the survey were registered by road side number plate registration. Subsequent telephone interviews and mail-back customized stated preference questionnaires were then used. The revealed preference data was obtained from the same individuals. Their actual travel times were taken from the traffic simulator CONTRAM, calibrated for the Stockholm network. Data on actual travel time uncertainty was obtained from traffic cameras.

2. Model framework

In trip timing modelling, scheduling disutility is usually formulated as departure time shifts from the most preferred departure time, PDT. Different variables represent time shift earlier, SDE (schedule delay early), and later, SDL (schedule delay late), defined as

$$\mathrm{SDE}_{in} = \max[\mathrm{PDT}_n - \mathrm{DT}_{in}, 0] \tag{1}$$

$$\mathrm{SDL}_{in} = \max[\mathrm{DT}_{in} - \mathrm{PDT}_n, 0] \tag{2}$$

where DT_{in} is the departure time of alternative i and individual n. If defining scheduling delay with respect to arrival time, SDE and SDL are instead based on the shift between preferred arrival time PAT_n and arrival time AT_{in}. PDT is, in the present study, defined as the departure time the driver would choose if there are never any queues on the road network. PAT is defined as PDT plus the travel time the driver would face if there are no queues on the road network.

[1] A set of express lanes on an otherwise free and congested road offers high-quality service to people who are willing to pay a time-varying toll.