1 INTRODUCTION

In circumstances where panel type data is available on choices made by individuals it is reasonable to assume that previous experiences somehow condition these choices. In the context of this presentation, the choices are routes that car drivers take to travel through the City of York, United Kingdom. The route choice information is available through a vehicle registration plate survey and the explanatory data through journey times, distances and the nature of the traffic infrastructure along the routes. We demonstrate that there is circumstantial evidence to suggest that there is a habit effect in the choices individuals make and suggest ways to incorporate this effect in choice models. This will lead to a better understanding of the uncertainty in urban dynamic route choice.

2 SURVEY DATA

The data used in this study arises from a number plate matching exercise conducted in the City of York, United Kingdom, during the Summer and early Autumn of 2000. The survey was a 100% survey, conducted during a one hour period of the morning peak, 08:00 to 09:00. The actual survey dates were 27th, 28th June, 7th, 8th, 9th, 11th, 13th, 27th September and 18th October, 2000. A map of the survey points is shown in figure 1. The 11th and 13th September were during a week of fuel protests across the UK, when there was a general perception that petrol would be in short supply due to blockades of oil refineries. In addition, Lendal Bridge (survey point H) was unavailable to private traffic after 10th September. The original survey data may be found at the project web site, hosted at the University of York.

Clark and Clegg (2001) provide a report on the impact of the fuel protests and the bridge closure. One element of this analysis is a discrete choice logit model of drivers’ route choice, as a function of journey time, distance, infrastructure and impact dummy variables. Whilst their
model produces reasonable and plausible estimates for the parameters in the model, they treat their data as a cross-sectional pooled dataset and fail to take account of its time series nature.

3 TIME DIMENSION

Route data from this same data set suggests that of those vehicles which are seen on two adjacent weekdays, over 50% are seen to travel the same route on both days. As the time span between the two days increases, this percentage drops, so for example, two survey days 14 working days apart, the percentage is nearer 35% to 40%.

Re-examining the data in more detail and concentrating on one origin-destination pair, namely AJE or ACE, and the (almost) four “consecutive” days of 7th, 8th, 11th and 13th September, number of features are apparent. Firstly the shorter distance route, ACE, is chosen on 60 of the 92 journeys recorded in this dataset. The journey times, however, show that route AJE is marginally the quicker of the two routes, taking 7:50min to travel 2.1m, whilst ACE takes 8:30min to travel a shorter 1.8km.

Turning our attention to the repeating nature of the journeys, one sees that three vehicles were seen to travel between this origin-destination on all four days, eight on any three consecutive
days and the remainder, 28 vehicles, on any two consecutive days. Of these 39 repeat vehicles, only on three occasions (out of the 56 possibilities), did they follow a different route on a subsequent day. This suggests that the data set contains a high degree of habitual information in how vehicles (and hence drivers) select which route to follow. This habit may be attributable to different effects for different drivers, e.g. lack of knowledge of the road network; conservatism of the driver or a constraint due to family or other commitments.

If explicit account is taken of this habitual behaviour within a modelling framework, then its strength can be estimated and measured. More seriously, failure to account for this degree of repetition and route experience may undermine the validity of any models estimated on this data set. In particular, these time period correlations may bias the other, cross-sectional, parameters in the model. These are our justifications for using multi-period choice models.

4 PREVIOUS STUDIES

The first significant piece of work that took onboard the concept of incorporating previous experience in choice models was by Daganzo and Sheffi (1979). A summary of the model may also be found in Ortùzar and Willumsen (1990), but basically the approach involves the incorporation of lagged utilities within the current period’s utility expression. This model of a discrete choice time-series and state-dependence model is, however, only an approximation to a true state-dependence relationship. For an initial application of this model form to surveyed transport data, see Johnson and Hensher (1982), where seven different models were fitted to mode choice data. The models ranged from a simple model that took no account of the panel nature of the data, through dummy variable models and onto full experience models. The value of in-vehicle time saving was shown to vary considerably amongst the models (from 23 to 204 cents per person), highlighting the need for a correct consideration of the panel nature of the data.

A consideration of the analysis of multi-period panel data is provided by Chamberlain (1984) in the context of fixed and random effects linear models. This work is described and extended in Hsiao (1986) which presents two chapters on how to model discrete panel data, both for complete and incomplete datasets, using the linear probability, logit or probit approach. The approach adopted allows for both static and dynamic relationships to exist. A number of papers in a volume edited by Hsiao et al (1999) also provide the theory and demonstrate the application of autoregressive or dynamic models to discrete panel data. Within this collection,
most of the applications are, however, based on simulated data and none of the real world data sets are from a transport context.

Other approaches to these datasets include a multi-period multinomial probit model, introduced by Louviere and Hensher (2000), which allows for the incorporation of individual effects and a flexible variance co-variance structure. A fully formed state dependent model is described in Heckman (1981). This model formulation has the ability to incorporate explanatory factors, past histories, cumulative experience and habit persistence. The formulation is, however, rather complex to estimate.

5 FURTHER STUDY

The presentation will describe an analysis of this data set that will take explicit account of the day-to-day dynamics or habits in individual vehicle route choice. Much of the analysis will build on the work described in the previous section, using the formulations outlined. Estimates will be made of the strength of the day-to-day relationship and how taking account of such effects affects the other parameters in the model. In particular, the implications and utility of the results for planners and policy makers will be covered.

Initial effort has been directed towards the estimation of multiperiod logit and probit formulations using the LIMDEP (version 7.0) software package. As of writing, there remain some issues outstanding on how the “autocorrelation” parameter is estimated within the software and we are actively pursuing our investigations into these issues. It is also our intention to use code written in the powerful GAUSS (version 3.6) matrix/econometric language that will allow for a greater degree of flexibility in the formulation of the model. Two approaches under consideration are a simulated maximum likelihood/moments estimation approach (Keane, 1994 & Geweke, Keane and Runkle, 1994) and a conditional maximum likelihood estimation approach (Chamberlain, 1984). The GAUSS code offers the facility to specify AR(1) errors and individual specific random effects for each choice, thereby identifying the dynamic features within the data.

There will also be an opportunity to further investigate these dynamic changes using a follow on data set of vehicle registration plates collected in York during the Autumn of 2001.
REFERENCES


