Transportation Research Part A 106 (2017) 51–64

Contents lists available at ScienceDirect

Transportation Research Part A

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

Measurement of non-random attrition effects on mobility rates using trip diaries data

Lissy La Paix Puello[a,*], Marie-José Olde-Kalter[b], Karst T. Geurs[a]

[a] Centre for Transport Studies, Faculty of Engineering Technology, University of Twente, Drienerlolaan 5, 7500AE Enschede, The Netherlands
[b] Goudappel Coffeng, Snipperlingsdijk 4, 7417 BJ Deventer, The Netherlands

ARTICLE INFO

Keywords:
Panel data
Hybrid choice models
Mobility rates
Attrition
Completeness

ABSTRACT

This paper examines the influence of panel attrition on the intrapersonal dynamics in self-reported trip rates, using the data from the 2013, 2014 and 2015 waves of the Netherlands Mobility Panel, a large scale household panel. A hybrid choice model (HCM) was developed to simultaneously model the effect of socioeconomic, infrastructure and land use variables, life events and non-random attrition on trip rates, whereby the latent variable (LV) model is composed of panel attrition and survey completeness. The discrete choice model (DCM) includes four trip rate categories, including zero trips. The probability of each trip rate category was estimated for both the HCM and the DCM models; with and without the LV model. The first main conclusion from this paper is that the largest bias due to panel attrition occurs in the probability of reporting no trips per day, and 1–2 trips per day. Also, the HCM models show a correlation between the probability of reporting no trips per day and the tendency to drop out altogether. The second main conclusion is that the results show that the latent variables (attrition and completeness) are statistically significant in estimating mobility. Also, socioeconomic variables (gender, driving license, household type and size), mode preferences, spatial infrastructure and life events determine mobility rates and remain significant after adding attrition/completeness variables. Thirdly, the results proved that attrition effects significantly vary across waves.

1. Introduction

In most countries, the understanding of people's travel behaviour is based on cross-sectional travel surveys in which only one day is surveyed for each respondent; often also in 'representative' periods when traffic flows are maximal (see for an overview Ortúzar et al., 2011). This is not enough to gain a proper understanding of the dynamics in travel behaviour and the behavioural changes needed to reverse the worrying long-term trends of growing mobility, congestion, increasing oil consumption and greenhouse gas (GHG) emissions (Ortúzar et al., 2011). More specifically, cross-section travel surveys do not give any information to ascertain how choices will vary over time (i.e. policy response) if the system changes. Moreover, from the onset, the travel demand models that form the basis of transport policy making in many countries have been based upon these one-day cross-section surveys (Stopher and Zhang, 2011). The models implicitly assume that behaviour is adjusted to new circumstances instantaneously (i.e. behaviour is assumed to be in equilibrium, fully adjusting to the prevailing values of contributing factors) and travel patterns are highly repetitive in the short run. Both assumptions, however, do not hold. In the literature it is often reported that there are all sorts of factors that will not lead to travellers immediately adapting to new situations (e.g., see for overviews of literature Fujii, 2010; Meurs, 2007). Furthermore, there

* Corresponding author at: Centre for Transport Studies, Faculty of Engineering Technology, University of Twente, Drienerlolaan 5, 7500AE Enschede, The Netherlands.
E-mail addresses: l.c.lapaixpuello@utwente.nl (L. La Paix Puello), MOldeKalter@goudappel.nl (M.-J. Olde-Kalter), k.t.geurs@utwente.nl (K.T. Geurs).

http://dx.doi.org/10.1016/j.tra.2017.09.002
Received 28 November 2016; Received in revised form 21 August 2017; Accepted 1 September 2017

is a strong need for a larger data collection for survey data, which would also allow, for example, drawing a significant origin-destination matrix for traffic modelling (Castaigne et al., 2009).

Panel surveys trace the same individuals over time, and offer unique opportunities to both analyse and model the dynamics of travel behaviour. Such modelling can be performed at the individual level and enable identification of inter- and intrapersonal variation on travel behaviour. However, from the literature, it is well known that multi-day panel studies can introduce a bias. This means that respondents' abilities and intentions to keep accurate travel diaries, for a certain period of time, decreases. This bias is linked to the fact that trip diary data are collected from the same respondents at two or more points within the period (Kitamura and Bovy (1987), Ortúzar et al. (2011)). If respondents participate in a number of waves of a panel survey, fatigue can occur and introduce non-random variations in reported trips. Therefore, drop-outs are a common characteristic of all panel studies, occurring when certain sample units leave the panel in the second or subsequent waves. Annual and biannual transport panels typically lose between 20 and 40 percent of their participants per wave (Polak, 1999). For some panel experiments, attrition does not only occur *between* waves, but also *within* waves. Attrition *within* means that respondents do not complete part of the survey, within the same wave. On the other hand, attrition *between* means that respondents drop out from one wave to the next.

This paper aims to model intra- and interpersonal dynamics in trip rates while controlling for possible attrition biases in the panel data, using data from the Netherlands Mobility Panel (in Dutch: MobiliteitsPanel Nederland – MPN) (Hoogendoorn-Lanser et al., 2015). A hybrid choice model is developed to simultaneously model trip rates using socioeconomic and other travel behaviour variables and non-random attrition.

To the authors' knowledge, this is the first simultaneous model for non-random attrition and travel behaviour. Plus, we use data from the Netherlands Mobility Panel, currently the largest general purpose mobility panel in the world.

2. Literature review: travel behaviour and panel attrition

Understanding the dynamics in sociodemographic and attitudinal factors can well be called a crucial task in the analysis of mobility rates. There is evidence that at least 50% of the improvement in the overall model statistics can be due to the presence of repeated observations (Cherchi and Cirilo, 2008), and there is a strong element of habit or persistence in certain travel behaviour elements, e.g. household car ownership, from one year to the next (Nolan, 2010). Even so, Stopher and Zhang (2011) have shown that there is relatively little intrapersonal variation, which in their case referred to repetition of tours from one day to the next. It denotes the relevance of collecting high quality longitudinal data. The more variable behaviour is, the more flexible the supply needs should be.

A proper estimation of frequency of trips is the base for estimating travel behaviour. For example, frequency of trips is the best explanation for mode choice habits (Cherchi et al., 2013). Also, it has been found that several elements play an important role to estimate trip rates. For example, changes in lifestyle (Ma and Goulias, 1997; Meurs et al., 1989), changes in neighbourhood (Meurs and Haaijer, 2001), life-events, such as changing jobs or moving house (Clark et al., 2016), preferences for specific transport modes such as bicycle (Van Wee et al., 2002) and purpose-mode preferences (Schwanen and Mokhtarian, 2005), and use of the internet for shopping or work, within the household and/or individually (Farag et al., 2006; Francke and Visser, 2015; Kenyon, 2010). Therefore, a main advantage of panel data is to show individual day-to-day, but also year-to-year variation in terms of lifestyle and life events. However, a high risk point of panel data is *attrition effects*, this one and other relevant factors are elaborated in the following section.

2.1. How to measure non-random attrition and its correlation with trip rates?

Panel surveys are identified in the literature as an alternative to cross-section data collection (e.g., see Kitamura, 1990; Ortúzar et al., 2011; Zumkeller and Ottmann, 2009). Well-researched long duration panels are the 1971 Uppsala Household survey covering a 5-week time period (e.g., Hanson and Huff, 1988; Huff and Hanson, 1986), the six-week Mobi*drive* survey for the German cities of Karlsruhe and Halle, and related to this a survey in Thurgau (Switzerland). The Mobi*drive* panels have been used for example for analysis of the rhythms of daily life (Axhausen et al., 2002), and to measure variability in travel behaviour (Schlich and Axhausen, 2003).

Most panels reported in the literature are unrepeated short duration surveys due to the modest respondent burden and costs, compared to repeated and long duration panels. 'Unrepeated' means that multiple days travel data is collected but the survey is not repeated, or not with the same respondents. Furthermore, most repeated short duration panels in the literature were designed for specific purposes or projects. The Santiago panel, for example, was designed as a repeated (5 wave) short duration panel to evaluate the effects of the introduction of the Transantiago public transport system (Yáñez et al., 2010a). In Germany and the Netherlands general purpose mobility panels are used. The German Mobility Panel (MOP) has been conducted annually since 1994, using a seven-day trip diary (Zumkeller and Chlond, 2009), but it uses a rotating panel sample. The Mobility Panel for the Netherlands (MPN), a three-day trip diary, has been conducted annually since 2013 (Hoogendoorn-Lanser et al., 2015); and this is the largest ongoing panel in the world repeated with the same respondents.

However, there is a long-standing discussion over the reliability of longitudinal data in showing the real mobility patterns and *attrition* effects. Literature shows that *attrition* is almost always non-random; that is, the units that drop out are systematically different from the units that remain. For instance, research has shown that attrition can be related to households with lower incomes, educational levels, occupational status and less active mode use (see, for example, Kitamura and Bovy (1987) Pendyala et al. (1993),Hensher et al. (1992) and Brownstone and Chu (1997)). Lifestyles, too, might influence the *completeness* of the survey. For example, more mobile respondents can be more reluctant to complete the survey.