Transportation Research Part A 67 (2014) 110–126

Contents lists available at ScienceDirect

Transportation Research Part A

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

Using repeated cross-sectional travel surveys to enhance forecasting robustness: Accounting for changing mode preferences

Khandker M. Nurul Habib [a,*], Joffre Swait [b], Sarah Salem [a]

[a] Department of Civil Engineering, University of Toronto, Canada
[b] Institute for Choice, University of South Australia, Australia

ARTICLE INFO

Article history:
Received 7 June 2013
Received in revised form 9 June 2014
Accepted 11 June 2014

Keywords:
Mode choice
Model transferability
Context influences
Preference stability
Work trip

ABSTRACT

Transportation system capacity and performance, urban form and socio-demographics define the influences and constraints conditioning the preferences of urban residents for different transport modes. Changes in characteristics of urban areas are likely to lead to changes in preferences for alternative modes of transport over time; as a consequence, statistical models to forecast mode choice need to be sensitive to both purposeful changes to urban systems as well as exogenous shocks. We make use of the 1996, 2001 and 2006 household surveys conducted in the Greater Toronto and Hamilton Area to study mode preference evolution and model forecasting performance. These repeated cross-sectional household surveys provide an opportunity to investigate aggregate structural changes in commuting mode preferences over time, in a manner sensitive to changes in the urban area. We focus on commuting mode choices because these trips are prime determinants of peak period congestion and peak spreading. We then address how to combine the three cross-sections econometrically in a robust way that allows for use of a single mode choice model across the entire period. Using independent data from 2012, we are able to compare the individual year and combined models in terms of forecasting performance to demonstrate the combined model's more robust forecasting performance into the future.

1. Background and motivation

Transportation supply (transportation system capacity and performance), urban form and socio-economic characteristics help define the opportunities, resources and constraints operating on mode preferences of urban residents. Policy implementation and changes over time in policy emphasis affect investments in the transportation system, which subsequently impact transportation system performance as well as urban form, and less directly, socio-demographics. This is observable for the Greater Toronto and Hamilton Area (GTHA), which has experienced significant variation in policy focus over the last three decades, from auto- to transit-oriented policies, to which are added the effects of smart growth land use policies. Despite efforts to reduce automobile and encourage transit use, peak period traffic congestion in GTHA nonetheless continues to increase.

Besides the obvious effects that characteristics of the social, economic, and transport systems have on mode choice, it is possible that policy-instigated changes to the transport system will not only create impacts on users via changes in transport

* Corresponding author.
E-mail addresses: khandker.nurulhabib@utoronto.ca (K.M.N. Habib), Joffre.Swait@uts.edu.au (J. Swait), sarah.salem@mail.utoronto.ca (S. Salem).

http://dx.doi.org/10.1016/j.tra.2014.06.004
0965-8564/

system performance (cost, time, accessibility, etc.), but will also affect commuting mode choice decision difficulty. That is to say, transport system changes may generate decision contexts in which it is more difficult for decision makers to correctly identify the optimal (i.e. highest utility) modal alternative, ultimately leading to a higher incidence of sub-optimal mode choices across the population, with adverse implications not only for individual travelers but for transport system performance. Swait and Adamowicz (2001) have suggested that decision difficulty can be recognized in Random Utility Models by appropriate parameterization of scale factors in choice models.

Most commuting mode choice models presented in the literature are based on cross-sectional datasets of travel diary surveys (*inter alia*, recent examples are Heinen et al., 2012; Tsamboulas et al., 2012; Habib, 2012; Zaman and Habib, 2011). In addition to modal characteristics, such cross-sectional sources can also be used to investigate the effects of various contemporaneous contextual factors involved in such situations (e.g., Long et al., 2010). However, it is possible that preferences for commuting mode choice evolve over time, but mode choice models developed using only one cross-sectional travel survey will by definition not be able to capture such changes. Clearly, if mode choice preferences are actually changing but mode choice models used to forecast future demand do not allow for these changes, forecasts will necessarily be biased.

One way to deal with changing preferences is to abandon cross-sectional travel surveys and use instead panel data, i.e., data based on tracking behavior of the same individuals over an extended time period (see, e.g., Golob et al., 1997). Panel data allows tracking changing behavior consistently over time. While panel data is the most suitable sources, however, panel data is unlikely to be available to most transportation planning agencies for any number of organizational, logistical and financial reasons. Arguably, in terms of the needs dictated by long-term planning functions in these agencies, panel data may even be contra-indicated. Hence, these agencies' planning missions are still strongly structured around cross-sectional data (in which sample composition is unique to each sampling effort) collected at intervals of five, ten, or even more years (Miller, 2007). In the Toronto, Canada, region, for example, household surveys have been conducted every 5 years since 1986 (DMG, 2012); in Sydney, Australia, every 10 years from 1971, then continuously from 1997 (Inbakaran and Kroen, 2011); Chicago did a survey in 1990, then only repeated a major household regional data collection nearly 20 years later in 2007/2008 (Inbakaran and Kroen, 2011). Despite this kind of wide variation in survey interval practices and known difficulties (e.g. cost, response rates, under-reporting) in the collection of point-in-time household survey data, as well as its inability to represent time-based influences on behavior, we join Stopher and Greaves (2007) in surmising that household surveys will be with us for some time to come. Moreover, we also concur to the fact that whenever available, panel data should be used solely or even in conjunction of crossectional datasets to investigate changed in preferences.

The question we pose, then, is how to account for preference evolution in the development of long-range forecasting models under these historical, contextual and institutional constraints? If panel data are not available and if a single cross-sectional dataset cannot by definition contain preference evolution information, then the logical answer that presents itself is that we must make use of multiple cross-sections (at least) collected at different points of time in the region of interest. Obviously, the changes in transportation system between two consecutive surveys importantly affect the choice patterns. If there is no significant change happened in the systems in between two crossectional surveys, both dataset would result similar model either independently or in pooled form. However, more than two crossectional dataset would, perhaps, help capturing the immediate effects their easing pattern over time of any abrupt changes between two surveys. Availability of any small span or small sample panel dataset in conjunction of multiple repeated crossectional datasets will be even more beneficial.

This paper therefore proposes a specific procedure and econometric formulation for a mode choice model that is sensitive to aggregate changes in preferences over time, but makes use of multiple traditional cross-sectional travel surveys. We use three such repeated cross-sectional household travel surveys collected from the same study area over the period of 10 years to detect whether changes occurred over time in both systematic and random components of mode choice utilities, including linking utilities to individual-level contextual factors arising from transport system performance, urban form and socio-demographic characteristics. We find evidence of mode preference changes for the study area; we are then led to develop a combined time-sensitive mode choice model that forecasts to the three cross-sections essentially as well as year-specific models, increasing confidence that the combined model can be used for context-sensitive forecasting. We are also able to conduct an external validity test by using the various models to predict to an independent data set from 2012, 6 years beyond the most recent survey data used in the estimation exercise.

With this overarching objective in mind, the paper is arranged as follows: the next section presents a brief literature review on the temporal transferability of cross-sectional mode choice models; this is followed by a section explaining the econometric model formulation used in this study; then a succinct description of the Toronto area datasets we utilize is presented, followed by a section presenting the empirical models and results on temporal transferability of year-specific models. These results motivate the penultimate section, which focuses on a combined cross-sections model that we believe represents a more robust tool for forecasting because of its context-sensitivity and description of mode preference evolution. The paper concludes with a discussion of key findings and an agenda for further research.

2. Literature review

At the heart of our research question (i.e., accounting for mode choice preference evolution using disaggregate cross-sectional data sets) is the underlying phenomenon of mode choice parameter stability over time within a given region of interest. As a testament to the importance of this issue, we draw upon the literature concerning the temporal transferability of