Does the social context help with understanding and predicting the choice of activity type and duration? An application of the Multiple Discrete-Continuous Nested Extreme Value model to activity diary data

Chiara Calastri [a], Stephane Hess [a,*], Andrew Daly [a], Juan Antonio Carrasco [b]

[a] Institute for Transport Studies and Choice Modelling Centre, University of Leeds, UK
[b] Universidad de Concepción, Chile

A B S T R A C T

An understanding of activity choices and duration is a key requirement for better policy making, in transport and beyond. Previous studies have failed to make the important link with individuals' social context. In this paper, the Multiple Discrete-Continuous Nested Extreme Value (MDCNEV) model is applied to the choice of activity type and duration over the course of two days, using data from the Chilean city of Concepción. In common with other studies, heterogeneity across decision makers is accommodated in the model by analysing the impact of different socio-demographic, mobility and residential location variables on both the activity choice and the time allocation decision. In addition, different social network and social capital measures are found to be significantly correlated with the choice and duration of different activities, and we show how these relationships seem to differ from the effects of socio-demographic variables. Finally, we perform a forecasting exercise using the MDCNEV model, highlighting the differences in substitution patterns from a standard MDCEV model.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

In recent years, activity-oriented approaches have gained considerable ground in the study of travel behaviour (Axhausen and Gärling, 1992). Travel demand is believed to be mainly a derived demand, directed at objectives such as going to work or performing recreational activities (Bhat et al., 2013; Ettema and Timmermans, 2003). The understanding of activity scheduling, which includes the decision of which specific activities to perform, with whom, for how long and using which transportation mode (Doherty et al., 2002; Gärling et al., 1998), can in turn lead to greater insights into the drivers of travel behaviour. Initial contributions to the literature treated the different dimensions of activity choice (such as type, timing and duration) separately, while in the last decade a growing amount of literature has highlighted the value of jointly investigating these aspects (Bowman and Ben-Akiva, 2001; Ettema et al., 2007).

The first econometric models accommodating both the discrete and continuous dimensions of choice were developed starting from the late 1950s by Tobin (1958), Heckman (1977), Dubin and McFadden (1984), Train (1986) and De Jong

---

* Corresponding author.
  E-mail address: s.hess@its.leeds.ac.uk (S. Hess).

(1990). Starting by using a system of equations, each corresponding to one choice dimension (Bhat, 2001), Chandra Bhat and his co-authors gradually developed a more general and flexible framework to model the choice of multiple alternatives and a continuous amount associated with each of them, in the form of the Multiple Discrete-Continuous Extreme Value (MDCEV) model (Bhat, 2008). This model has been applied in several studies analysing activity choice and duration (e.g. Bhat, 2005; Bhat et al., 2006; Kapur and Bhat, 2007) and constitutes the state of the art in modelling multiple discrete-continuous choices. These studies concluded that socio-demographic characteristics of individuals and households, ownership and availability of mobility tools, accessibility and land use characteristics are significant determinants of the choice of the different activities. For example, Kapur and Bhat (2007) study weekend day activity engagement by participants in the 2004 American Time Use Survey. Their results show how low income households are more likely to perform in-home activities, such as in-home leisure or maintenance, a conclusion that reflects the financial constraints to performing the generally more expensive out-of-home activities. Individual socio-demographics also affect activity choice. Women are for example more likely to be involved in household maintenance, with the same applying to married as opposed to singlerespondents, while middle-age people (40–60) are found to be more likely to engage in arts and events. A limited number of studies applied the nested version of the model (MDCNEV) to study time allocation (Pinjari and Bhat, 2010b; Rajagopalan et al., 2009; Bernardo et al., 2015).

While most datasets collect information about respondents' socio-demographic characteristics alongside time use diaries, contextual information about the circumstances in which people make their choices is often not available. However, there is clear scope for a relationship between an individual's social environment and his/her travel and activity choices. We are careful here in not positing a specific directionality of this relationship at the outset. Indeed, if a relationship between a large social network and the choice of out-of-home recreation is found, then it may of course be tempting to infer that the person conducts many such activities as a result of having many friends, However, it is similarly possible that the person developed a large social network to facilitate him/her performing out-of-home recreational activities. While the challenges in terms of causality remain, it clearly still important to test for these effects in models to understand the relationship between the choices and the context in which they are made. This is one of the aims of the present paper. At the same time, it is also important to test for confounding between these contextual variables and other socio-demographic characteristics, another point we pay careful attention to.

The closest existing work has got to this issue has come in attempts to find an impact of the social dimension on leisure and social activities, specifically on their frequency (Carrasco and Miller, 2009) and duration (van den Berg et al., 2012), while some work has also jointly modelled several dimensions (Moore et al., 2013; Carrasco and Habib, 2009; Habib et al., 2008). These efforts showed the importance of considering the social dimension to explain engagement in social and leisure activities, highlighting the relevance of the cultural context examined (Kowald et al., 2013). One of the aims of the present work is to investigate the broader relations of the social dimension with time use going beyond just leisure and social activities, by looking at the time allocation for entire days.

The remainder of this paper is organised as follows. The next section describes the data used for our analysis, followed by a discussion of the modelling framework. We then present our application and the different models we estimated. After describing our results, we forecast with the MDCEV and the MDCNEV models and discuss the implications of including the social network variables for model performance and forecasting. We conclude by drawing policy considerations and suggesting directions for future work.

2. Data

2.1. Survey and data collection

The dataset used for our analysis was collected in 2012 within the *Communities in Concepción* project, which involved people from four neighbourhoods of the Chilean city of Concepción. Concepción is located approximately 500 km south of the capital Santiago and with its 1 million population it constitutes the second largest urban centre of the country. Two of the neighbourhoods (*Agüita de la Perdiz* and *La Virgen*) are close to the city centre, with the first one being a medium–high income neighbourhood, and the second being a medium–low income one. The other two (*S. Sabina* and *Lomas S. Sebastian*) are further away from the city. Medium–low income households mainly populate the first one, while the second one is home to medium-high income people. The specific sampling approach adapted for this study implies that there is not enough variability to control for accessibility, walkability and other measures normally used to describe the built environment characteristics.

The data were collected face-to-face in respondents' homes. Participants were initially asked to complete a detailed socio-demographic questionnaire, including questions about themselves, their family composition and their mobility and communication tool ownership. They were then asked to complete a 2-day activity diary by filling a grid with detailed information about the activities they have been engaged in during one recent weekday and one recent weekend day, and during which time slots these took place.

In addition to these more traditional components, respondents were asked to elicit their social network by completing a so-called "name generator". This technique, extensively used in the sociology (Campbell and Lee, 1991) and travel behaviour (Carrasco et al., 2008, Kowald et al., 2010, Pike, 2014) literature, consists of asking people to recall their entire social network